

PF 02-APR-1987; 87JP-0076285.
 XX
 XX 02-APR-1986; 86JP-0076285.
 XX
 XX (RIKA) RIKAGAKU KENKYUSHO.
 XX
 XX WPI; 1987-325245/46.
 DR N-PSDB; AAN70651.
 XX
 XX DNA sequence coding cellulase gene - is derived from *Bacillus* sp. no.
 PT 1139Z and can hydrolyse cello-triose and cello-tetraose.
 XX
 XX Disclosure: Fig 3-1 - 3-3; 11pp; Japanese.
 PS
 XX This cellulase hydrolyses cellobiose or cello-tetraose and does not
 CC hydrolyse cellobiose. Molecular weight is approx. 92k dalton and
 CC optimum pH is 9.
 CC
 XX Sequence 800 AA:
 SQ

Query Match 4.6%; Score 108.5; DB 8; Length 800;
 Best Local Similarity 19.9%; Pred. No. 0.26; Indels 173; Gaps 25;
 Matches 96; Conservative 54; Mismatches 159;

QY 8 LRLSLPCFALIAIOQAQVNPVAFVDEVRSEND---LGQDNELPIDVQASQTASDTPT 64
 DB 10 LLSLIIIVLIIISI-----fptalaegntrednfkhllgndnvkrpseagaldqdvvg 64
 QY 65 ANPL-DEHEPELY-----TTALENKTMLINCSAL-----NODIWRILACY-DTLVHGFT 110
 DB 65 qmtlvdqhgkqlrgmsthgqlwfpelndnaykalandwesnmrlamygengyaasn 124
 QY 111 PAVIKTKRSIRIDETI-----WQT-----IKGKPOVY 137
 DB 123 PELKSRVTKGIDLAENDMYIVDWHVHAPGPRDPRVYAGAEFFRIDIAALYMPNHII 184
 QY 138 YQETTPD-----IFLMGNEKMLTKKDAKOLEYAKQFTPLSFSFDRN-----NTPU 186
 DB 185 YELANEPSSNMNGGAGIPNNEEGWNAVK-----EYADPIVEMLRDSGNADHIIIVGSPN 239
 QY 187 WSSRP-----HNPMYVLPPIF--MHGKPNRS--PMTPSHEAKQFTPNFRAPAEIKF 232
 DB 240 WQGRPLAENDPIDDHHTMYTVHFYGSAASTESYPPETPNSEGNVMSNTRYA-----1 295
 QY 233 QVSVVYKKAEDLMGT-----DSDLWFGYTOQ-----SHWQIFNGKNS-----R 270
 DB 236 ENGVAVFATE--WYLSQANGDGARYFDEADVLEFINENIISWAMSLTN--KNEVSGAFT 352
 QY 271 PPRVH-----DYQPELFLTPQVYSDLPW--- 293
 DB 353 PFEIGKSNATSIDPPDQVWPEEISLSEGYVYRARIKYNYEP---IDTKYTKYVWFDN 409
 QY 294 DCKVRKMGAVHNSGES-----AKLSRSNRRAYLMAGMEWKNLTYVPRIM 340
 DB 410 dgtkgqfgv-----ngdsrpedvvleneagalklsgldasndvsegunwanarisadgw 463
 QY 341 GR 342
 DB 464 gk 465

RESULT 15
 ID AAR26021
 ID AAR26021 standard: Protein; 822 AA.

XX AAR26021;
 AC
 XX 02-FEB-1993 (first entry)
 DT
 XX Alkaline cellulase K-64.
 DE
 XX Mass production; recombinant.
 KW

XX
 OS *Bacillus* sp. KSM-64 (FERM P-10482).
 XX
 XX JP04190793-A.
 PN
 XX 09-JUL-1992.
 PD
 XX 26-NOV-1990; 90JP-0324401.
 PF
 XX 26-NOV-1990; 90JP-0324401.
 PR
 XX (KAOS) KAO CORP.
 PA
 XX WPI; 1992-280112/34.
 DR N-PSDB; AAG27180.
 XX
 XX Alkaline cellulase mass produ. - using recombinant plasmid and
 PT microorganism contg. cellulase gene
 PS
 XX Disclosure: Fig 4; 14pp; Japanese.
 CC
 XX Alkaline cellulase K-64 may be mass produced recombinantly in host
 CC microorganisms such as *E. coli* HB101, C600 or JM109, *B. subtilis* BD170,
 CC 168 or ISW1214.
 XX
 XX Sequence 822 AA:
 SQ

Query Match 4.6%; Score 108.5; DB 13; Length 822;
 Best Local Similarity 19.9%; Pred. No. 0.27;
 Matches 96; Conservative 54; Mismatches 159; Indels 173; Gaps 25;

QY 8 LRLSLPCFALIAIOQAQVNPVAFVDEVRSEND---LGQDNELPIDVQASQTASDTPT 64
 DB 10 LLSLIIIVLIIISI-----fptalaegntrednfkhllgndnvkrpseagaldqdvvg 64
 QY 65 ANPL-DEHEPELY-----TTALENKTMLINCSAL-----NODIWRILACY-DTLVHGFT 110
 DB 65 qmtlvdqhgkqlrgmsthgqlwfpelndnaykalandwesnmrlamygengyaasn 124
 QY 111 PAVIKTKRSIRIDETI-----WQT-----IKGKPOVY 137
 DB 123 PELKSRVTKGIDLAENDMYIVDWHVHAPGPRDPRVYAGAEFFRIDIAALYMPNHII 184
 QY 138 YQETTPD-----IFLMGNEKMLTKKDAKOLEYAKQFTPLSFSFDRN-----NTPU 186
 DB 185 YELANEPSSNMNGGAGIPNNEEGWNAVK-----EYADPIVEMLRDSGNADHIIIVGSPN 239
 QY 187 WSSRP-----HNPMYVLPPIF--MHGKPNRS--PMTPSHEAKQFTPNFRAPAEIKF 232
 DB 240 WQGRPLAENDPIDDHHTMYTVHFYGSAASTESYPPETPNSEGNVMSNTRYA-----1 295
 QY 233 QVSVVYKKAEDLMGT-----DSDLWFGYTOQ-----SHWQIFNGKNS-----R 270
 DB 236 ENGVAVFATE--WYLSQANGDGARYFDEADVLEFINENIISWAMSLTN--KNEVSGAFT 352
 QY 271 PPRVH-----DYQPELFLTPQVYSDLPW--- 293
 DB 353 PFEIGKSNATSIDPPDQVWPEEISLSEGYVYRARIKYNYEP---IDTKYTKYVWFDN 409
 QY 294 DCKVRKMGAVHNSGES-----AKLSRSNRRAYLMAGMEWKNLTYVPRIM 340
 DB 410 dgtkgqfgv-----ngdsrpedvvleneagalklsgldasndvsegunwanarisadgw 463
 QY 341 GR 342
 DB 464 gk 465

Search completed: November 30, 2001, 14:17:02
 Job time: 384 sec

CC sequence of H. pylori (ATCC 55679) was determined from overlapping
 CC contigs generated by mechanically shearing the bacterial DNA. The
 CC sequences were analysed for ORF of at least 180 nucleotides, and the
 CC predicted coding regions defined by computer evaluation. To identify
 CC likely H. pylori antigens for vaccine development, the amino acid
 CC sequences predicted from various ORF were analysed for significant
 CC homology to other known or exported membrane proteins. Having identified
 CC and determined the sequences of interest, particular regions can be
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide
 CC production, e.g. in E. coli hosts.

XX Sequence 253 AA;

Query Match 6.6%; Score 156.5; DB 18; Length 253;
 Best Local Similarity 25.6%; Pred. No. 1e-06; Mismatches 72; Indels 65; Gaps 8;
 Matches 56; Conservative 26;

QY 157 KDAKOLEVAAKQFTPLSLSFDDLDRNTPLMSSRPHNPMYVLPFMHGKPRSPNTPSHE 216
 Db 72 kkylnmddygytyfllpfyhsf-----tpflfyghnpn----- 105
 QY 217 AKQFTPEFAPRLKFOVSVKAAEDLWGTDSDFMGYTOQSHWQJFNGKNSRPRVHD 276
 Db 106 ---ygrnef-----kfgistfivpfirhlwtkgtlylayqtwtfglyndpqasapmrm 157
 QY 277 YQPEIFLTQVYSDLPMDGKV---RMIGMGAVHHSNG-ESAKLSRSNRAVLMAGMEMKN 332
 Db 158 fmpellyvyp1-nktrpdygkignfseilwqhlisngvgaqcyqpin----- 204
 QY 333 LTWMPRIWGRIFREGSGSQ--PDDNPDLIDYGYGDVRF 369
 Db 205 -----kegnpenqfpgqprvlyvkdyngqkdvtw 231

RESULT 13

AAW70991
 ID AAW70991 standard; Protein; 1686 AA.

XX AAW70991;

DT 19-OCT-1998 (first entry)

XX Human class II P13 kinase-C2alpha.

XX Human: class II phosphoinositide 1-phosphatase; P13 kinase;

KW PIK3-C2alpha; class II; resistance; Wortmannin; LY294002.

XX Homo sapiens.

OS

PN W09832864-A2.

PD 30-JUL-1998.

PF 27-JAN-1998; 98W0-GB00244.

PR 28-JAN-1997; 97GB-0001652.

XX (LUDM-) LUDMIG INST CANCER RES.

PI Domin J, Waterfield MD;

DR WPI: 1998-427960/36.

DR N-PSDB; AAV42920.

XX New nucleic acid encoding phosphoinositol kinase 3-C2 alpha or its

PT fragments - useful for, e.g. treatment of tumour cells where

PS phenotype is associated with expression of kinase

XX Claim 3; Fig 1; 52pp; English.

CC The present sequence represents a human class II phosphoinositide 1-phosphatase (P13) kinase designated PIK3-C2alpha. It is characterised as a class II

CC kinase due to the presence of a conserved C2 domain found in murine and
 CC Drosophila class II P13 kinases, its apparent lack of a p85 binding site
 CC and a substrate affinity to inositol lipids p85 and p85(4)P. The
 CC protein has resistance to P13 kinase inhibitors Wortmannin and LY294002.
 CC Antibodies against the protein (optionally humanised), are used to
 CC identify class II P13 kinases. Antisense sequences, antibodies or
 CC dominant negative mutants of the P13-C2alpha protein, are useful in human
 CC or veterinary medicine to block class II kinases. They can be used to
 CC treat tumour cells where the phenotype is associated with expression of
 CC P13-C2alpha protein.

XX Sequence 1686 AA;

Query Match 5.0%; Score 117; DB 19; Length 1686;
 Best Local Similarity 19.2%; Pred. No. 0.13; Mismatches 154; Indels 140; Gaps 21;
 Matches 86; Conservative 67;

QY 45 QDNELPIDV-----QSATQASATDTPANPLDE-----HEPELYTVALENTMLIN----- 88
 Db 522 eddetpvdlnkhlgyiekpckeamtlrhpveellasyhmqvelaqienqhravqylkav 1581
 QY 89 ---CSALNODIMRLACYDTLVHGETPAVYKTKRSIRLDETWTQTKGKPOVYQETDPI 1145
 Db 582 rkicsald-gyeltal-----tesvykkikravnlpr-----ktadv 618
 QY 146 FLWNGENKMLTKKAKOLEVAAKQFTPLSLSFDDLDRNTPLMSSRPHNPMYV-----LP 199
 Db 619 slfge-----darsst-rgslnpenpavgslnqlta 651
 QY 200 IF-----MHGKPRSPNTPSHEAKQFTPEFAPRLKFOVSVKAAEDL---WGTDSDFL 252
 Db 652 lydlrlrhansgrprtdaqskvkewtteqlgflf-----aahglisnwnsnyeky 707
 QY 253 FGYYQOSHWOJFNGKN-SRPF---RVHDIQPEIFLTQ-----PV-YSDLPMDGKYRMI 300
 Db 708 ylicslsh---ngkdlfkprskkvgtyknfilylkwdellfpdlsqpllesvhl 763
 QY 301 GMGAVHHSNGSEAKLSRSN-----RAYLMAGMEMKNLTWMPRIWGRIFK 345
 Db 764 lfgllngssgspsdnkqkgspealgyvlpdclfrflltcg-----ckillyw----- 812
 QY 346 EGSSGPDNDPDIIDYGYGDVRFYOLENKSNIQTVRNPBSGKALQDYYVPLGK 405
 Db 813 --tshsnsvpgrvttkgymeritlvqdfspafdllytppqdrslilqnmlettend 870
 QY 406 ISGYFQIFQGYGOSLIDYNEATSFGV 432
 Db 871 lkg-----klidlhkdsisl 887

RESULT 14

AAW70420
 ID AAW70420 standard; Protein; 800 AA.

XX AAW70420;

DT 20-JAN-1991 (first entry)

XX Sequence encoded by cellulase gene derived from Bacillus sp. No. 1139.

XX Enzyme; cellotriose; cellotetraose; hydrolysis.

XX Bacillus sp. No. 1139.

OS

PN JF62232386-A.

PD 12-OCT-1987.

XX Key Location/Qualifiers

XX Peptide 1..30

XX Protein 31..800

QY 259 SHWQIFN-GKNSRPFVRVHDYQPEIFLTQPVYSDLPMWDGKVRMIGMGAHVHHSNGESAKLSR 3177

```

QY 157 KDAKOLEYAAKQFTPLSLFDLDRNNTPLMSSRPHNPMYVLPINHGKPNPSPTPSHE 210
      ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 69 kkylnmmdyigtlyflpfyhsf-----tpifqyghnrlp----- 102

```


XX	01-MAY-1998;	98US-0083756.
PR	31-JUL-1998;	98US-0094859.
PR	02-SEP-1998;	98US-0098994.
PR	02-SEP-1998;	98US-0099062.
PR	09-OCT-1998;	98US-0103749.
PR	09-OCT-1998;	98US-0103794.
PR	09-OCT-1998;	98US-0103796.
PR	25-FEB-1999;	99US-0121528.
XX	(CHIR) CHIRON CORP.	
PA	(GENO-) INST GENOMIC RES.	
XX		
PI	Fraser C, Galeotti C, Grandi G,	Hickey E, Masignani V, Mora M;
PI	Petersen J, Pizza M, Rappunli R,	Ratcl G, Scariato E, Scarselli M;
PI	Tettelein H, Ventor JC;	
XX		
DR	WPI: 2000-062150/05.	
XX	N-PSDB; AAZ53920.	
PT		
PT	Novel Neisserial polypeptides predicted to be useful antigens for	
XX	vaccines and diagnostics -	
PS		
XX	Claim 2; Page 905; 1453pp; English.	
CC	AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAZ75941	
CC	represent novel <i>Neisseria meningitidis</i> and <i>N. gonorrhoeae</i> polynucleotides	
CC	and polypeptides. AAZ54537 to AAZ54576 and AAZ54615 to AAZ5473 represent	
CC	PCR primers used in the exemplification of the present invention. The	
CC	polypeptides, the polynucleotides, antibodies and compositions of	
CC	the invention can be used as vaccines, as diagnostic reagents, and as	
CC	immunogenic compositions. The polypeptides can be used in the	
CC	manufacture of medicaments for treating or preventing infection due to	
CC	Neisserial bacteria (e.g. meningitis and septicemia), to detect the	
CC	presence of <i>Neisseria</i> bacteria, or to raise antibodies. They may also	
CC	have been used to screen for agonists or antagonists, which may themselves	
CC	may also be used in gene therapy protocols.	
XX		
XQ	Sequence 370 AA;	

Query Match	34.5%;	Score 815;	DB 21;	Length 370;
Best Local Similarity	44.7%;	Pred. No. 7.5e-70;		
Matches 163;	Conservative 67;	Mismatches 111;	Indels 24;	Gaps 8;
QY	87	INCSALNODIRKLACYPDLVHGFEPPAVI-----KTKRSIRLDETIYQTI-KGKPOVYOE	140	
Db	21	lqcaalctdnvrclaeaydrifaagqlpsasgqegqekavlnltetrrssldkgeaivavek	80	
QY	141	TTDPITFLMGNEKGMITTKDKAOLEYAAKOFTPLSLFSDLDNRNN-TPRLSSRPHPMVIPL	199	
Db	81	ggdal-----padsageaaiaivrlslmydlckldrglllgvrehnmpmylmp	127	
QY	200	IFMHCKPKRSPPTPSH-EAKOFTPEHFRAPELKPOVSUKVAEELMGSDSLMGYQO	258	
Db	128	lwymnspfnayapsprcrgttvgekfsgqqkraekikqvstkskaeeldtkrraalwgyqqr	187	
QY	259	SHMOIEN-GKNSRPPRVHDYOPELFLTOPVYSDLPMDCKVRMIGMAVHNSGESAKLSR	317	
Db	188	sdwqynggrksapfrnctdykpeflctgpkvadrlfggrltmlgsgfvbhqnsqgrpsr	247	
QY	318	SMNRATYLMAGKEKKMLUTMPRIWGIIFKEGSSSQDDNDPDLIDYGYGQDVRLYOLEKMS	377	
Db	248	swnlrtaymagnewgkllvrlpvrwvaiafq-sgdk-ndnppdladmgnygdvkvkyrlndrq	305	
QY	378	NISGTVRYNPPSSGKALDLDVYVPLGKISIGYFOIGYQOSLDIYNEEATSFYGLMLN	437	
Db	306	nvyaslrvnptkygataaaytfrpklyklygvvrgfhyggesldiyhnhkqngigtlnfn	365	
QY	438	DMWGL 442		
Db	366	dlgdgl 370		

RESULT	8
AAV70629	.
ID	AAV70629 standard; Protein; 374 AA

DT 18-JUL-2000 (first entry)

DE *Neisseria meningitidis* serogroup B strain H44/76 BASB033 protein.

KW BASB033; diagnosis, prophylaxis; treatment, antibacterial; vaccine;
KW *Neisseria meningitidis* infection.

OS *Neisseria meningitidis*.

PN W0200015801-A1.

PD 23-MAR-2000.

PF 09-SEP-1999; 99WO-EP06718.

PR 14-SEP-1998; 98GB-0020003.

PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

PI Ruelle J;

DR WPI; 2000-271439/23.

XX

PT meningitidis, useful for diagnosis, prophylaxis and treatment of N.

XX XX

XX
XX

CC Neisseria meningitidis serogroup B strain H44/76. The protein

CC phospholipase A. The present sequence is useful for diagnosis,

CC used for the discovery and development of antibacterial compounds and

XX

50 Sequence 374 AA;

Query Match	34.58; Score 815; DB 21; Length 374;
-------------	--------------------------------------

Matches 163; Conservative 67; Mismatches 111; Indels 24; Gaps 8;

87 INCSALNODIMRLACYDTLVHGETPAVI-----KTKRSIRLDETIWQTI-KGKPQVVYQE 140

Db 25 lqcaaltndvtrlacyrdrifaaqlpssagqegqeska1nltetvrssldkgeavivvek 84

141 TTDPILMGNEKMLTKKDAKQLEYAAKQFTPLSLSEDLDRNN-TPLWSSRPHPMPVLP 199

Db 85 ggda1-----padsagetadlytprlsmydidkndlrgl1lgvrehnmpy1mp 131

QY 200 IFMHGKPNRSPNTPSH-EAKQFTPNEFRAPELKQVSVKKAEDLWGTDSDLWFGYTQO 258

Db 132 lwynspnyapgsptrgltvqekfgqkraetklqvsfksklaedlffltradlwfgytlqr 191

QY 259 SHWQIFN-GKNSRPFVRVHDYQPEIFLTQPVYSDDLPWDGKVRMIGMGAVHHNGESAKLSR 317

Db 192 sdwqlynggrksapfrntdykpeifltqpvkadlpfggrlrmigagfvhqsngsrsrpestr 251

QY 318 SWNRAYLMAGMEWKNLTVMPRIWGRIFKEGSGSQPDDNPDLIDYGYGDVRFLYOLENKS 377

Db 252 swnrlyamagmgkltvlpvrfa^{dd}-sgdk-ndnp^ladymgygdvklqyrln^{dr}q 309

QY 378 NISGTVRYNPRSGKALQLDVYPLGKGISGYFOIFQGYGSLIDYNHEATSFVGMLN 437

```
Db      21  lqcaaltdnvtlacydrifaaglpssaaqgeqgeskavlnltetvrsldkgeavlyvek 80
QY      141  TTDPIFLMGNEKGLTKKDAKOLEYAAKOFPLSLSFDLDBNN-TPLWSSRPHNPMVLP 199
Db      81  ggda1-----padsagetadlycpislmvldkndrlgllgvrehnmpmlymp 127
QY      200  IFMHGKPNRSPNTPSH-EAKQFTPNFERAPBLKFOVSVKVAADLWGTSDLMFGYTOQ 258
Db      128  fwynspnyapsptgtvtgqekfgqkkractklyqsfkskiaeelnlktadlwfgytr 187
QY      259  SHMOIFN-GKNSRPFRRNDYQPELFLTPQVYSDLPMDGKVMGMGVHNSGSAALS 317
Db      188  sdwqlyngyrksapfrntdykpeflftqpvkadlprfgirlmlygaftvhsngsqrspre 247
QY      318  SMNRAYLMAGMEKNLTVMPRIMGRIFKEGSGSDPDNDPILDVYGYGDVRFVLEENKS 377
Db      248  swmrlyamagmewgklvlprrvvratfdq--sgdk--ndnpdladymgvgdvklyqrlndrq 305
QY      378  NISGTVRYNPRSGKALQLDVYVPLGKISGYFOIFOGYGOSLIDYNHAEATSFVGMLN 437
Db      306  nvysvlrnpktygalaayaftfplkgklkgvvrfgfhgyselldynhknqngiglmfn 365
QY      438  DMWGL 442
Db      366  dwdgl 370

RESULT 6
AA75157
ID  AA75157 standard; Protein; 370 AA.
AC  AA75157;
DT  21-MAR-2000 (first entry)
DE  Neisseria meningitidis ORF 582 protein sequence SEQ ID NO:1788.
XX
DE  Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW  antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW  antibacterial; gene therapy.
XX
OS  Neisseria meningitidis.
XX
PN  WO957280-A2.
XX
PD  11-NOV-1999.
XX
PF  30-APR-1999; 99WO-US09346.
XX
PR  01-MAY-1998; 98US-0083758.
PR  31-JUL-1998; 98US-0094869.
PR  02-SEP-1998; 98US-0098894.
PR  02-SEP-1998; 98US-0099062.
PR  09-OCT-1998; 98US-0103749.
PR  09-OCT-1998; 98US-0103794.
PR  09-OCT-1998; 98US-0103796.
PR  25-FEB-1999; 99US-0121528.
XX
PA  (CHIR ) CHIRON CORP.
PA  (GENO-) INST GENOMIC RES.
XX
PI  Fraser C, Galicotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI  Petersen J, Pizzo JC, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI  Tettelin H, Venter UC;
XX
DR  WPI; 2000-062150/05.
DR  N-PSDB; AA253919.
XX
XX  Novel Neisserial polypeptides predicted to be useful antigens for
PT  vaccines and diagnostics
PS  Claim 2; Page 904; 1453pp; English.
```

```
XX      AA253015 to AA254536, AA254577 to AA254615, and AA74253 to AA75941
CC  represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
CC  and polypeptides. AA254537 to AA254576 and AA254616 to AA254731 represent
CC  PCR primers used in the exemplification of the present invention. The
CC  polypeptides, the polynucleotides, antibodies and compositions of
CC  the invention can be used as vaccines, as diagnostic reagents, and as
CC  immunogenic compositions. The polypeptides can be used in the
CC  manufacture of medicaments for treating or preventing infection due to
CC  Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC  presence of Neisseria bacteria, or to raise antibodies. They may also
CC  be used to screen for agonists or antagonists, which may themselves
CC  have use as antibacterial agents. The polynucleotides of the invention
CC  may also be used in gene therapy protocols.
XX
SQ      Sequence 370 AA;

Query Match 34.5%; Score 815; DB 21; Length 370;
Best Local Similarity 44.7%; Pred. No. 7.5e-70;
Matches 163; Conservative 67; Mismatches 111; Indels 24; Gaps 8;

QY      87  INCSALMODIMRLACYDTLVHGETRPAYI-----KTKRSIRLDETINQTI-KGKQVYYOE 140
Db      21  lqcaaltdnvtlacydrifaaglpssaaqgeqgeskavlnltetvrsldkgeavlyvek 80
QY      141  TTDPIFLMGNEKGLTKKDAKOLEYAAKOFPLSLSFDLDBNN-TPLWSSRPHNPMVLP 199
Db      81  ggda1-----padsagetadlycpislmvldkndrlgllgvrehnmpmlymp 127
QY      200  IFMHGKPNRSPNTPSH-EAKQFTPNFERAPBLKFOVSVKVAADLWGTSDLMFGYTOQ 258
Db      128  fwynspnyapsptgtvtgqekfgqkkractklyqsfkskiaeelnlktadlwfgytr 187
QY      259  SHMOIFN-GKNSRPFRRNDYQPELFLTPQVYSDLPMDGKVMGMGVHNSGSAALS 317
Db      188  sdwqlyngyrksapfrntdykpeflftqpvkadlprfgirlmlygaftvhsngsqrspre 247
QY      318  SMNRAYLMAGMEKNLTVMPRIMGRIFKEGSGSDPDNDPILDVYGYGDVRFVLEENKS 377
Db      248  swmrlyamagmewgklvlprrvvratfdq--sgdk--ndnpdladymgvgdvklyqrlndrq 305
QY      378  NISGTVRYNPRSGKALQLDVYVPLGKISGYFOIFOGYGOSLIDYNHAEATSFVGMLN 437
Db      306  nvysvlrnpktygalaayaftfplkgklkgvvrfgfhgyselldynhknqngiglmfn 365
QY      438  DMWGL 442
Db      366  dldgl 370

RESULT 7
AA75158
ID  AA75158 standard; Protein; 370 AA.
AC  AA75158;
DT  21-MAR-2000 (first entry)
DE  Neisseria meningitidis ORF 582 protein sequence SEQ ID NO:1790.
XX
DE  Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW  antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW  antibacterial; gene therapy.
XX
OS  Neisseria meningitidis.
XX
PN  WO957280-A2.
XX
PD  11-NOV-1999.
XX
PF  30-APR-1999; 99WO-US09346.
XX
```


QY 181 RNNTPLMSSRPHPNMYVLPFIMHGKPNRSPNTPSHEAKQFTPNNEFRAPELKFOYSVKVA 240
 CC and to identify agonists and antagonists. The polypeptides, antibodies,
 CC agonists and antagonists (which are bacteriostatic) are used for the
 CC treatment and prevention of diseases such as otitis media in infants and
 CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
 CC invasive diseases, and chronic otitis media with hearing loss. The
 CC polypeptides, agonists and antagonists are also used for screening of
 CC antibacterial drugs. The BASB034 products of the invention can be used
 CC screen for new antibacterial compounds that may target resistant
 CC bacteria.

QY 241 AEDLMGTDSDLMRGYTOQOSHMOJFNGKNSRPFVNDYQPEFLITQPYVSLPMDGKVRMT 300
 CC
 Db 241 aedltwgtcdslwfygtqgsfmwqifngknsrpfvndyqpeflitqpyvslpmdgkvrmt 300
 CC
 QY 301 GMGAVHHSNGESAKLSRSNRRAYLMAGMEWKNLTVMPRIGRIFFKESGSGOPDNDPILD 360
 CC
 Db 301 gmgavhhsngesaklsrswrraylmagmewknlvmpriigrifksgsgqpdndpild 360
 CC
 QY 361 YYGXGVDFRFLYQLENKSNISGTVRNPRSGKALQLDVYVPLGKISGYFOIFQGYGQSL 420
 CC
 Db 361 yygvgdvrflyqlenksnlsqvtvyrprsgkqalqldvvyplgklsyfqifqgygqsl 420
 CC
 QY 421 IDYNHEATSFYGVGLMDNMGL 442
 CC
 Db 421 idynheatsfygvlmdnmgl 442
 CC

RESULT 3
 AAY85270
 ID AAY85270 standard; Protein: 442 AA.
 XX
 AC AAY85270;
 XX
 DT 29-JUN-2000 (first entry)
 XX
 DE BASB034 amino acid sequence #3.
 XX
 XX Moraxella catarhalis infection; BASB034; diagnosis; staging;
 KM vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
 KM sinusitis; nosocomial infection; invasive disease; chronic otitis media;
 KM hearing loss; antibacterial drug.
 XX
 OS Moraxella catarhalis.
 XX
 PN MO200015802-A1.
 XX
 PD 23-MAR-2000.
 XX
 PE 14-SEP-1999; 99WO-EP06781.
 XX
 PR 14-SEP-1999; 98GB-0020002.
 XX
 PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX
 PI Ruelle J;
 XX
 DR WPI: 2000-271440/23.
 XX
 DR N-PSDB: AAA10702.
 XX
 PT Novel BASB034 polynucleotides and polypeptides from Moraxella
 XX catarhalis used to prepare vaccines against bacterial infections
 PS Claim 3: Page 68; 106pp; English.

This sequence represents a Moraxella catarhalis BASB034 polypeptide from
 strain Mc2913. The invention relates to BASB034 polypeptides from
 M. catarhalis strains Mc2913, Mc2908, Mc2913 and Mc2969. The BASB034
 polynucleotides and polypeptides may be employed as research reagents and
 material for the discovery of treatments and diagnostics for diseases,
 particularly human diseases. They are particularly used to diagnose and
 treat M. catarhalis infections. They can be used for diagnosis of
 disease, staging of disease, or determining response of an infectious
 organism to drugs. The polynucleotides may be used as a source for
 hybridization probes, and for screening of genetic mutations, serotype,
 organism or strain identification, identification of mutations in BASB034
 sequences, and as components of arrays which are useful for diagnostic
 and prognostic purposes. The polypeptides can be used to produce
 antibodies. The polypeptides can also be used in vaccine formulations,

CC and to identify agonists and antagonists. The polypeptides, antibodies,
 CC agonists and antagonists (which are bacteriostatic) are used for the
 CC treatment and prevention of diseases such as otitis media in infants and
 CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
 CC invasive diseases, and chronic otitis media with hearing loss. The
 CC polypeptides, agonists and antagonists are also used for screening of
 CC antibacterial drugs. The BASB034 products of the invention can be used
 CC screen for new antibacterial compounds that may target resistant
 CC bacteria.

QY 1 MKVSLSTFLTSLTFCFAILAIQQAAPNPVAFVDEVRESENDLGQNELPIDVQATQSA 60
 CC
 Db 1 mkvslstflstlscfalaiaiqgakvnpvafvdevrsendlgqnelpidvqatqsa 60
 CC
 QY 61 STDANPLDDEHEPRLTYTALENKTMNLNCSALNODIMRLACYDPTLVGEPVATKTRST 120
 CC
 Db 61 stdanpldehepelytalaenktmlncsalnqdlmrlyacydptlvgepvatktkrs1 120
 CC
 QY 121 RLDETIWQTIKGRPOVYVOETTDPIFLMGNEKGMLTRKDAKOLEYAKQFTPLSLSPDL 180
 CC
 Db 121 rldetiwtikgrpqvvyvoettdpiflmgnekgmtlrrkdaqlkleyaakqftplsld 180
 CC
 QY 181 RNNTPLMSSRPHPNMYVLPFIMHGKPNRSPNTPSHEAKQFTPNNEFRAPELKFOYSVKVA 240
 CC
 Db 181 rnntpmwssrphnmyvlpfimgkpnrsptpsheakqftpnnefrapeltfgyvsvkva 240
 CC
 QY 241 AEDLMGTDSDLMRGYTOQOSHMOJFNGKNSRPFVNDYQPEFLITQPYVSLPMDGKVRMT 300
 CC
 Db 241 aedltwgtcdslwfygtqgsfmwqifngknsrpfvndyqpeflitqpyvslpmdgkvrmt 300
 CC
 QY 301 GMGAVHHSNGESAKLSRSNRRAYLMAGMEWKNLTVMPRIGRIFFKESGSGOPDNDPILD 360
 CC
 Db 301 gmgavhhsngesaklsrswrraylmagmewknlvmpriigrifksgsgqpdndpild 360
 CC
 QY 361 YYGXGVDFRFLYQLENKSNISGTVRNPRSGKALQLDVYVPLGKISGYFOIFQGYGQSL 420
 CC
 Db 361 yygvgdvrflyqlenksnlsqvtvyrprsgkqalqldvvyplgklsyfqifqgygqsl 420
 CC
 QY 421 IDYNHEATSFYGVGLMDNMGL 442
 CC
 Db 421 idynheatsfygvlmdnmgl 442
 CC

RESULT 4
 AAY85269
 ID AAY85269 standard; Protein: 442 AA.
 XX
 AC AAY85269;
 XX
 DT 29-JUN-2000 (first entry)
 XX
 DE BASB034 amino acid sequence #2.
 XX
 XX Moraxella catarhalis infection; BASB034; diagnosis; staging;
 KM vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
 KM sinusitis; nosocomial infection; invasive disease; chronic otitis media;
 KM hearing loss; antibacterial drug.
 XX
 OS Moraxella catarhalis.
 XX
 PN MO200015802-A1.
 XX
 PD 23-MAR-2000.
 XX
 PR 14-SEP-1999; 99WO-EP06781.

XX Claim 3; Page 69; 106pp; English.

XX This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
XX strain Mc2969. The invention relates to BASB034 polypeptides from
XX M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
XX polynucleotides and polypeptides may be employed as research reagents and
XX material for the discovery of treatments and diagnostics for diseases,
XX particularly human diseases. They are particularly used to diagnose and
XX treat M. catarrhalis infections. They can be used for diagnosis of
XX disease, staging of disease, or determining response of an infectious
XX organism to drugs. The polynucleotides may be used as a source for
XX hybridization probes, and for screening of genetic mutations, serotype,
XX organism or strain identification, identification of mutations in BASB034
XX sequences, and as components of arrays which are useful for diagnostic
XX and prognostic purposes. The polypeptides can be used to produce
XX antibodies. The polypeptides can also be used in vaccine formulations,
XX and to identify agonists and antagonists. The polypeptides, antibodies,
XX agonists and antagonists (which are bacteriostatic) are used for the
XX treatment and prevention of diseases such as otitis media in infants and
XX children, pneumonia in elderly, sinusitis, nosocomial infections and
XX invasive diseases, and chronic otitis media with hearing loss. The
XX polypeptides, agonists and antagonists are also used for screening of
XX antibacterial drugs. The BASB034 products of the invention can be used
XX screen for new antibacterial compounds that may target resistant
XX bacteria.

XX Sequence 442 AA:

Query Match 100.0%; Score 2363; DB 21; Length 442;
Best Local Similarity 100.0%; Pred. No. 9, 3e-219;
Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVSLSTLSTLPCFAITAIQQAQAVPVAFVDEVSENDLGDNDLPIDVOSATOSA 60
DB 1 mkvslstlslstlpcfallaiaqgaavpvaafvdevsendlgdndlpdvgsatgsa 60
QY 61 STDTPNPLDEHPELYTTALENKMILINCSALNODIMRLACDYTLVHGTPAVITKRSI 120
DB 61 stdtanpldehepelyttalenkmlincsalngdimrlacydlvhgtpaviktrsi 120
QY 121 RLDETIWOTIKGKPOVVOYQETTDPIFLMGNEKGMLTRKDAKOLEYAAKOFPLSLSPDLD 180
DB 121 rldetiwtikgkpvvyqettdpiflmgnekgmtlrrkdakoleyaakofplslsfdld 180
QY 181 RNNTPLMSSRPNNPVYLPFIMHGKRPNSPNTPSHBAKOFTPNEFRAPBLAKQVSVKKA 240
DB 181 rnntplwssrpnnpvylpifimhgkprnsntpsbhaakoftpnefrapelkfqsvskva 240
QY 241 AEDLGTSDLMFQYTOOSHOIFNGKNSRPVRVDYQPEIRLQPVYSDLPMDCKVRMI 300
DB 241 aedlgtstdlmfgytqoshoifngknsrprvrhdyqpeirfqpyvsdldpmdckvrm 300
QY 301 GMGAVHNSNGESAKLSRSNMNRAVYLMAGHEWKNLTYMPRIWGRIFRFGSGSQPDNDPDIJD 360
DB 301 gmgavhnsngesaklsrsnmnraylmagmewknltympriwgrifrgsgsqdndpdiid 360
QY 361 YYGVDVRFVLIOLEKNSNIGTVRYNPRSGALQDLYVYPLGKISGYFOFGYQOSTL 420
DB 361 yygvdvrfvlylqleknsgnigstvrynprrsgalqldvvyplgkisygyfgygysgl 420
QY 421 IDYNHEATSPFGVGLMNDMMGL 442
DB 421 idynheatsfgyglmndwmg 442

RESULT 2
AA85268
ID AAY85268 standard; Protein; 442 AA.
XX
AC AAY85268;
XX

DT 29-JUN-2000 (first entry)

XX BASB034 amino acid sequence #1.
DE
XX
XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW hearing loss; antibacterial drug.
XX
XX Moraxella catarrhalis.
OS
XX WO200015802-A1.
PN
XX 23-MAR-2000.
PD
XX 14-SEP-1999; 99WO-EP06781.
PF
XX 14-SEP-1998; 98GB-0020002.
PR
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
PA
XX Ruelle J;
PI
XX WPI; 2000-271440/23.
DR
XX N-PSDB; AAA10700.
XX

PT Novel BASB034 polynucleotides and polypeptides from Moraxella
PS catarrhalis used to prepare vaccines against bacterial infections -
XX
XX Claim 3; Fig 2; 106pp; English.

XX This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
XX strain Mc2931 (ATCC 43617). The invention relates to BASB034 polypeptides
XX from M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The
XX BASB034 polynucleotides and polypeptides may be employed as research
XX reagents and material for the discovery of treatments and diagnostics for
XX diseases, particularly human diseases. They are particularly used to
XX diagnose and treat M. catarrhalis infections. They can be used for
XX diagnosis of disease, staging of disease, or determining response of an
XX infectious organism to drugs. The polynucleotides may be used as a source
XX for hybridization probes, and for screening of genetic mutations,
XX serotype, organism or strain identification, identification of mutations
XX in BASB034 sequences, and as components of arrays which are useful for
XX diagnostic and prognostic purposes. The polypeptides can be used to
XX produce antibodies. The polypeptides can also be used in vaccine
XX formulations, and to identify agonists and antagonists. The polypeptides,
XX antibodies, agonists and antagonists (which are bacteriostatic) are used
XX for the treatment and prevention of diseases such as otitis media in
XX infants and children, pneumonia in elderly, sinusitis, nosocomial
XX infections and invasive diseases, and chronic otitis media with hearing
XX loss. The polypeptides, agonists and antagonists are also used for
XX screening of antibacterial drugs. The BASB034 products of the invention
XX can be used screen for new antibacterial compounds that may target
XX resistant bacteria.

XX Sequence 442 AA:

Query Match 99.6%; Score 2354; DB 21; Length 442;
Best Local Similarity 99.5%; Pred. No. 6, 8e-218;
Matches 440; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKVSLSTLSTLPCFAITAIQQAQAVPVAFVDEVSENDLGDNDLPIDVOSATOSA 60
DB 1 mkvslstlslstlpcfallaiaqgaavpvaafvdevsendlgdndlpdvgsatgsa 60
QY 61 STDTPNPLDEHPELYTTALENKMILINCSALNODIMRLACDYTLVHGTPAVITKRSI 120
DB 61 stdtanpldehepelyttalenkmlincsalngdimrlacydlvhgtpaviktrsi 120
QY 121 RLDETIWOTIKGKPOVVOYQETTDPIFLMGNEKGMLTRKDAKOLEYAAKOFPLSLSPDLD 180
DB 121 rldetiwtikgkpvvyqettdpiflmgnekgmtlrrkdakoleyaakofplslsfdld 180

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:17:00 ; Search time 72.04 Seconds

(without alignments)
454,475 Million cell updates/sec

Title: US-09-787-083-8

Perfect score: 2363
Sequence: 1 MKVSLSTLTLSILPCFAILA.....YNHEATSPGVGLMNDMMGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /SIDSB8/gcgdata/geneseq/AA1980.DAT:*
2: /SIDSB8/gcgdata/geneseq/AA1981.DAT:*
3: /SIDSB8/gcgdata/geneseq/AA1982.DAT:*
4: /SIDSB8/gcgdata/geneseq/AA1983.DAT:*
5: /SIDSB8/gcgdata/geneseq/AA1984.DAT:*
6: /SIDSB8/gcgdata/geneseq/AA1985.DAT:*
7: /SIDSB8/gcgdata/geneseq/AA1986.DAT:*
8: /SIDSB8/gcgdata/geneseq/AA1987.DAT:*
9: /SIDSB8/gcgdata/geneseq/AA1988.DAT:*
10: /SIDSB8/gcgdata/geneseq/AA1989.DAT:*
11: /SIDSB8/gcgdata/geneseq/AA1990.DAT:*
12: /SIDSB8/gcgdata/geneseq/AA1991.DAT:*
13: /SIDSB8/gcgdata/geneseq/AA1992.DAT:*
14: /SIDSB8/gcgdata/geneseq/AA1993.DAT:*
15: /SIDSB8/gcgdata/geneseq/AA1994.DAT:*
16: /SIDSB8/gcgdata/geneseq/AA1995.DAT:*
17: /SIDSB8/gcgdata/geneseq/AA1996.DAT:*
18: /SIDSB8/gcgdata/geneseq/AA1997.DAT:*
19: /SIDSB8/gcgdata/geneseq/AA1998.DAT:*
20: /SIDSB8/gcgdata/geneseq/AA1999.DAT:*
21: /SIDSB8/gcgdata/geneseq/AA2000.DAT:*
22: /SIDSB8/gcgdata/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2363	100.0	442	21	BA5B034 amino acid
2	2354	99.6	442	21	BA5B034 amino acid
3	2348	99.4	442	21	BA5B034 amino acid
4	2339	99.0	442	21	BA5B034 amino acid
5	822	34.8	370	21	BA5B034 amino acid
6	815	34.5	370	21	BA5B034 amino acid
7	815	34.5	370	21	BA5B034 amino acid
8	815	34.5	370	21	BA5B034 amino acid
9	812	34.4	375	21	BA5B034 amino acid
10	246.5	10.4	355	19	AAV98871
11	239.5	10.1	356	19	AAV10960

12	156.5	6.6	253	18	AAW20760
13	117	5.0	1686	19	AAW70991
14	108.5	4.6	800	8	AAV70420
15	108.5	4.6	822	13	AA26021
16	107.5	4.5	824	18	AA23180
17	107	4.5	157	18	AAW20358
18	106	4.5	798	21	AA40925
19	106	4.5	798	21	AA40925
20	106	4.5	798	21	AA40925
21	102.5	4.3	537	22	AA83944
22	100.5	4.3	1726	18	AAW38756
23	98.5	4.2	467	20	AAV13378
24	98.5	4.2	467	22	AA88401
25	98.5	4.2	467	22	AA88401
26	98.5	4.2	467	22	AA88401
27	95.5	4.0	502	22	AA880246
28	95.5	4.0	516	22	AA880246
29	95.5	4.0	516	22	AA880246
30	95	4.0	682	17	AAW04359
31	95	4.0	888	22	AA85616
32	94	4.0	372	22	AA670751
33	94	4.0	761	20	AAW99084
34	94	4.0	1227	22	AA681501
35	94	4.0	1717	22	AA620498
36	93.5	4.0	467	19	AAW5057
37	93	3.9	600	22	AA675092
38	92.5	3.9	857	13	AA629814
39	91	3.9	1627	16	AA667538
40	90	3.8	522	22	AA895513
41	90	3.8	560	11	AA806815
42	90	3.8	564	21	AA810457
43	90	3.8	564	21	AA81952
44	90	3.8	790	22	AA893045
45	90	3.8	841	20	AA24318

ALIGNMENTS

RESULT 1	
ID	AAV85271 standard; Protein: 442 AA.
AC	AAV85271;
XX	
DT	29-JUN-2000 (first entry)
XX	
DE	BA5B034 amino acid sequence #4.
XX	
KW	Moraxella catarrhalis infection; diagnosis; staging;
KW	vaccine; bacteriostatic; treatment; prevention; pneumonia;
KW	sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW	hearing loss; antibacterial drug.
OS	Moraxella catarrhalis.
XX	
PN	MO200015802-A1.
XX	
PD	23-MAR-2000.
XX	
PF	14-SEP-1999; 99WO-EP06781.
XX	
PR	14-SEP-1998; 98GB-0020002.
XX	
PA	(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX	
PI	Ruelle J;
XX	
DR	WPI: 2000-271440/23.
XX	
PT	N-PSDB; AAA10703.
XX	
PT	Novel BA5B034 polynucleotides and polypeptides from Moraxella catarrhalis used to prepare vaccines against bacterial infections

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:25:05 ; Search time 16.07 Seconds
(without alignments)
238.768 Million cell updates/sec

Title: US-09-787-083-6

Perfect score: 2360
Sequence: 1 MKVSLSTLTSLTSLSCFAILA.....YNHEATSGVGLMLNDMMGL 442

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 45707 seqs, 8681013 residues

Total number of hits satisfying chosen parameters: 45707

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	352.5	14.9	282	5	US-09-897-516-7051 Sequence 7051, Ap
2	114	4.8	1686	5	US-09-355-160D-2 Sequence 2, Appl
3	101	4.3	719	5	US-09-815-242-13423 Sequence 13423, A
4	101	4.3	2384	5	US-09-897-516-4692 Sequence 4692, Ap
5	97.5	4.1	1478	5	US-09-801-368-52 Sequence 52, Appl
6	97.5	4.1	1478	5	US-09-487-558-52 Sequence 52, Appl
7	94	4.0	761	5	US-09-463-488A-63 Sequence 63, Appl
8	91	3.9	2590	5	US-09-897-516-7203 Sequence 7203, Ap
9	88.5	3.8	742	5	US-09-815-242-10634 Sequence 10634, Ap
10	88.5	3.8	742	5	US-09-815-242-10634 Sequence 10634, Ap
11	86.5	3.7	678	5	US-09-897-516-7356 Sequence 7356, Ap
12	84.5	3.6	1021	1	PCT-US01-27760-801 Sequence 801, App
13	83.5	3.5	802	5	US-09-897-516-6909 Sequence 6909, App
14	82.5	3.5	595	5	US-09-897-516-5307 Sequence 5307, App
15	82.5	3.5	2765	5	US-09-800-198-72 Sequence 72, Appl
16	82	3.5	594	5	US-09-338-723A-2 Sequence 2, Appl
17	81.5	3.5	462	5	US-09-815-242-10991 Sequence 10991, A
18	80.5	3.4	616	5	US-09-760-446A-1267 Sequence 1267, App
19	80	3.4	237	5	US-09-897-516-7775 Sequence 7775, App
20	80	3.4	295	5	US-09-978-189-415 Sequence 415, App
21	80	3.4	295	5	US-09-978-192-415 Sequence 415, App
22	80	3.4	295	5	US-09-978-697-415 Sequence 415, App
23	80	3.4	295	5	US-09-978-824-415 Sequence 415, App
24	79.5	3.4	688	5	US-09-801-368-378 Sequence 378, App
25	79.5	3.4	688	5	US-09-487-558-378 Sequence 378, App
26	79	3.3	362	5	US-09-620-394B-3688 Sequence 3688, Ap
27	79	3.3	373	5	US-09-620-394B-3687 Sequence 3687, Ap

28	79	3.3	391	5	US-09-620-394B-3686 Sequence 3686, Ap
29	76.5	3.2	1015	5	US-09-897-516-4696 Sequence 4696, Ap
30	76.5	3.2	2764	5	US-09-800-198-68 Sequence 68, Appl
31	76	3.2	602	5	US-09-569-037-6 Sequence 6, Appl
32	75.5	3.2	508	5	US-09-976-594-377 Sequence 377, App
33	75.5	3.2	1381	5	US-09-922-600-16 Sequence 16, Appl
34	75	3.2	750	5	US-09-801-368-280 Sequence 280, App
35	75	3.2	750	5	US-09-487-558-280 Sequence 280, App
36	75	3.2	1338	5	US-09-848-909-35 Sequence 35, Appl
37	75	3.2	1737	5	US-09-800-198-71 Sequence 71, Appl
38	75	3.2	2724	5	US-09-800-198-13 Sequence 13, Appl
39	75	3.2	2733	5	US-09-800-198-8 Sequence 8, Appl
40	75	3.2	2973	5	US-09-897-516-4394 Sequence 4394, Ap
41	74.5	3.2	384	5	US-09-897-516-7960 Sequence 7960, Ap
42	74	3.1	641	5	US-09-897-516-7585 Sequence 7585, Ap
43	73.5	3.1	471	5	US-09-815-242-10682 Sequence 10682, A
44	73.5	3.1	879	5	US-09-659-287A-2 Sequence 2, Appl
45	73.5	3.1	898	5	US-09-445-096A-11 Sequence 11, Appl

ALIGNMENTS

```
RESULT 1
US-09-897-516-7051
; Sequence 7051, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215, 161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 7051
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.
; US-09-897-516-7051

Query Match 14.98; Score 352.5; DB: 5; Length 282;
Best Local Similarity 39.28; Pred. No. 3.2e-25;
Matches 85; Conservative 30; Mismatches 93; Indels 9; Gaps 5;

QY 226 RAPLKFQVSKVKAADLMDCTDLMFGYTOQSHOIFNCKNSRPPFVHDYQEIFLTQ 285
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 73 RKDEVHGVSEFGFPIWGIYGENSLGASTYQKSMWLSNKRSPRETETPEQPLIAM 132
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 286 PV-YSDLPWQDKVYIMGAVHNGESAKLSRSMNRYALMAGMKNLTVMPIWGRIF 344
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 133 LTVDELAWG--HLMEIEAGINHQSGRPESTRSMNRYALMOKGMQVYDLKPMWRLT 190
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 345 KEGSGSPDDNPDLIDYVGCDVAFYLQI-ENKSNISGTAVYINRSKGALQIDYVPLG 403
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 191 E--SNRRDNPDMNRKMGYRLKVGHYGNGVFTMGC--RYNWSGSGGELGWSYPLS 245
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 404 KGISGYFOIFQGYGOSLIDYVNHHEATSGVGLMLNDMM 440
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 246 KHVRFYQALFSGYGESMIDYFRQTRFGICVMLNDML 282
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 2
US-09-355-160D-2
; Sequence 2, Application US/09355160D
```

GENERAL INFORMATION:
APPLICANT: Domin, Jan
TITLE OF INVENTION: Novel Lipid Kinase
FILE REFERENCE: 2332-1-004
CURRENT APPLICATION NUMBER: US/09/355,160D
CURRENT FILING DATE: 1999-10-01
PRIOR APPLICATION NUMBER: PCT/G898/00244
PRIOR FILING DATE: 1998-01-27
PRIOR APPLICATION NUMBER: 9701652.1
PRIOR FILING DATE: 1997-01-28
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 1686
TYPE: PRT
ORGANISM: Homo sapiens
US-09-355-160D-2

Query Match 4.8%; Score 114; DB 5; Length 1686;
Best Local Similarity 19.0%; Pred. No. 0.036;
Matches 85; Conservative 68; Mismatches 154; Indels 140; Gaps 21;

QY 45 QDNELPIDV-----QSATQASTDTANPLDE-----HEPELYTALLENKTMLIN-----88
DB 522 EDEETPVLDNKLHYQIEKPEKEMTRHPEVLELDYSYHNOVELAQIENOHRAVDQYIKAV 581
QY 89 ---CSALNODIRLACYDTLVHGETPAVYIKTKRSTRIDEITMOTIKGPOVYQETDPI 145
DB 582 KRICSLD-GVETLAI-----TESVKKLRANVLPPS-----KTADVT 618
QY 146 FLMGNEKMLTKRKDAKQLEYAAKQETPLSLSPDLNNTPLMSRPHNPMY-----LP 199
DB 619 SLFGGE-----DTSKRSST-RGSLNENPNVQSYINQLTAA 651
QY 200 IF-----MGKPRKSPPTPSHEARQFTPNFRAPELKFQYVKKAAEDL-----WGTSDIM 252
DB 652 IYDLRLRLHANSRSPDCKQSSSKYKAWTTTEQLQFTIF-----AAHGISSMNVSVYEXY 707
QY 253 FCGYQOSHQIFNGKN-SRPF-----RVHDYQPELIFLQ-----PV-YSDLPMDGKVRMI 300
DB 708 YLICSLSH-----NGKDLFPKIOSKKYGTAKNFYLLKMBELIIFPQISQPLESVLHLT 763
QY 301 GKGAVHNSGESAKLSRSWN-----RAYLMAGNEMKMLTVMPRIMGRIRK 345
DB 764 LFGILNOSGSSPDNSKQKGPALGKVSLLPCLDFRFLTCG-----TKLVLW-----812
QY 346 EESSGQPDNPDLIDYGGVAFYQLLENKSNISCTVKNRPSGGAQLQDLVYVPLGK 405
DB 813 --TSSHTNSVPQTVTKKGYMERIVLQVDFPSPAFDIYTTPOVDRSIIQOHMLETLEND 870
QY 406 ISGYFOIFQYGGOSLIDYVHNEATSFQV 432
DB 871 IKG-----KLIDILHKDSSLGL 887

RESULT 3
US-09-815-242-13423
Sequence 13423, Application US/09815242
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13423
LENGTH: 719
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-815-242-13423

Query Match 4.3%; Score 101; DB 5; Length 719;
Best Local Similarity 20.2%; Pred. No. 0.16;
Matches 97; Conservative 71; Mismatches 191; Indels 122; Gaps 25;

QY 3 VSLSTLSTL-----LSCPAILAIQOAKVPPVAVFVDEVRSENDL 43
DB 13 LSTSLVLAIAIVLGGVFFYVSKAPSLSEKLVATTSSKLYDNKNQIADIGSEBRY 72
QY 44 -GDNELPIDVQSAOS-----ASDTANPLDEHEPELYTALLENKTMLINCSALN 93
DB 73 NAQANDIPDLVKAIAIYSIEDHREFDRGIDITILIGAFILNLSQNSLQGS-----SALT 126
QY 94 QDITMLACYDTLVHGETPAVYIKTKRSTRIDEITMOTIKGPOVYQET-----TDPYFLMG 149
DB 127 QOQIKLKYESTYSQDTIS-----RKAQEMLAIQLEOKATKQELIYYIYINKYMSN 178
QY 150 NEKGMLTK-----KKAQKE-----YAAKQETPLSLSPDLNNTPLMSRPP 191
DB 179 GNYGMQTAQNTYKGDINNLSLPQALLAGMPQAPNPQDIPRSHPEAAQDRRLVYL--SEM 236
QY 192 HNPVYVLPFEMCKPNSPPTPSHEARQFTPNFRAPE-----LKTQVS-VKYKAAEDLW 245
DB 237 KNGGYI-----SAQYKQAVWPTITDGLQSLKSASNPAYMDNYLKEVINQVEETGYNLL 292
QY 246 GTSDIMFGYTG--QSH-WQIFNGKNSRPPRVHDYQPELIFLQPVYSDLPMDCKVRMIGM 302
DB 293 TTGMQVYTNVDQEAOKHLMIDYMTDEVVAYPPDDELQ-----VASTIVDVS-NGKV-IAOL 345
QY 303 GAVHNSN-----GESAKLSRSNNRAVLMAGNEMKMLTVMPRIMGRIRK 355
DB 346 GAHQSNSVSFGINQAVETNRDM-----GSTMKPITDYAPALEYGVY-ESTATYVHDE 397
QY 356 PDILDYVG-----YGDVRELYOLENKSNIQVRYN-----PRSGGALOLD 397
DB 398 P--YNNPGTTPYVNMNDRGFGFNITLQYALQGRNNVPVETLKKVGLNRAKKTFLNGLGID 455
QY 398 Y 398
DB 456 Y 456

RESULT 4
US-09-897-516-4692
Sequence 4692, Application US/09897516
GENERAL INFORMATION:
APPLICANT: Corbin, David R.
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Huesling, Joseph E.
APPLICANT: Krasomil-Osterfeld, Karina C.


```
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spitdonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215, 161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 4692
; LENGTH: 2384
; TYPE: PRF
; ORGANISM: Xenorhabdus sp.
; US-09-897-516-4692
```

```
Query Match 4.3%; Score 101; DB 5; Length 2384;
Best Local Similarity 20.8%; Pred. No. 0.93;
Matches 97; Conservative 58; Mismatches 124; Indels 188; Gaps 29;
```

```
QY 30 PVAFVDEVRNS-----ENDLGODNDELPIDVOSATOSASTDTANPLDEHEPELYTTA 79
DB 1015 PENFIDPTRLRKTKAFNDLENDLGO----- 1040
QY 80 LENKTMILNCSALNODIMRLACYDTLVHGFTPAVITKRSIRLDETITMOTIKGKPOVYQ 139
DB 1041 -----SRLNMDLQFAILRYLTERE-----RISNL--QLISG-----YI 1072
QY 140 ETTDPFILNGNEKMLTKDAKOLEYAAKOFPLSLSFDRNN--TPL-WS-----S 189
DB 1073 DGDGP---ISDRHYFIKGNNSPEVEY---YWRITLDISL-RDSNNIISPLAMSEKETTLS 1125
QY 190 RPHNPYVLPIFNHGRPN---RSPNTPSHEANQFTPNEFRABELKFOVSVKYKAEDWL 245
DB 1126 LSGTLAIPIPIVIGROYAIWVERATPLMGADQ-KPTDYRAINVKF---TEKOSNGEW 1180
QY 246 GTDSDLM--FGYQOSHMOIFNGKNRPRFVHD-----YQPIFILOPVY--SDLP 292
DB 1181 SAPNELRLDGDNDANCEYPTPKDK---RIPDNEHYLKDKEYKPKALIMAVDQREEDP 1335
QY 293 WDQKVMIGMAVHNHNGESAKLSRSM--NRAYIMAGMEKNITVMPR-----I 339
DB 1236 W-----MGVLLYDTKE--KDSSTWVKNKDY---LELRDLVLVDKTLIDDERKIVET 1282
QY 340 WGRIFKEGSGSQPDNDPILDIYGYGDVRF--LYOLENKS-----ISGTVR 384
DB 1283 WYKLF-----NNPDLQHHYAGTEKEFVYKIEETKKNHFNFLKPPRSPILTLIDVR 1333
QY 385 YNPRSGKALQLDYVYPLCKGISGYFOIFOGYOSLI--DYNHEATS 429
DB 1334 LDFTKTKLL-----MGR-----NSLIKFDNTNOAAS 1360
```

```
RESULT 5
US-09-801-368-52
; Sequence 52, Application US/09801368
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: Norman, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
```

```
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 52
; LENGTH: 1478
; TYPE: PRF
; ORGANISM: Saccharomyces cerevisiae
; US-09-801-368-52
```

```
Query Match 4.1%; Score 97.5; DB 5; Length 1478;
Best Local Similarity 23.0%; Pred. No. 0.97;
Matches 110; Conservative 50; Mismatches 175; Indels 143; Gaps 26;
```

```
QY 2 KVISSTILS-----ILSCFAILAIQAKAVNPVAVVDEVRSENDLGODNELPI---DV 53
DB 912 KVNRSNSTVSTNSIFYSPLKRGSKRYVSTSAAD--IFEENDTFADAPPPDSDD 970
QY 54 QSATOSASTDTANPLDEHEPELYTTALENK--TMLINCSALNODIMRLACYDTLVHGFTP 111
DB 971 SDDSSSDOLIWKKKTAPR---TNNEKKDEKSDNSHTSDELF---YDS---QTO 1019
QY 112 AVITKRSIRLDETITMOTIKGKPOVYQETDPIFLNGNEKMLTKDAKOLEYAAKOF 171
DB 1020 DKMERK-----MTFRSPPEVYQ-NLEKFPFPRANDLKPYT-----EGIASPTS 1061
QY 172 PLISFPLDRNNT-----PLMSSRPHNPVYLPIFNHG-----KPNRSPTP----- 213
DB 1062 PKSLSLSLSPNNAASSRTEBTPSRPVPPDSSYEFTODGLGKPKPLNQAATPRRTKTR 1121
QY 214 --SHEAQFTPNEFRABELKFOVSVKYKAEE-DLMTGSDLMFGYTOOSHMOIFNGKNSR 270
DB 1122 TIMEBASLARKN-----SVKLKROTKMKGT---RMVEVTEHMYHSIMKANNSK 1167
QY 271 PFRVHDYQPELFILOPVYSDLPWDGKVMIGMA-----VHNSNGESAKLSR----- 317
DB 1168 -----GEYKEFAW--MKGEMIGKGSFGAVYLCILNTVTGEMAVAKQVEVPKY 1211
QY 318 -SWNRATLMAGMEKNITVMPRIRNGRIFKEGSGSQPDNDPILDIYGYGDVRFILQLENK 376
DB 1212 SSOEATILS-----TV-----EALRSEYSTLKLDHLNITVOYLGF-----ENK 1249
QY 377 SNISGTVRNPYPRSGKALQLDYVYPLCKGISGYFOIFOGYOSLIDYNHEATSEFVG 434
DB 1250 NNIT-----SLFEYV--AGGSVGSILRMGTGRFDEPLI--KHLTYVLKGL 1291
```

```
RESULT 6
US-09-487-558-52
; Sequence 52, Application US/09487558
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: Norman, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/487,558
```

```

; CURRENT FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US/09/801,368
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 52
; LENGTH: 1478
; TYPE: prt
; ORGANISM: Saccharomyces cerevisiae
US-09-487-558-52

```

```

Query Match          4.1%; Score 97.5; DB 5; Length 1478;
Best Local Similarity 23.0%; Pred. No. 0.97;
Matches 110; Conservative 50; Mismatches 175; Indels 143; Gaps 26;

```

```

QY 2 KVSLSLTLSS-----ILSCFAILAIQAAKAVNPVAFVDEVSNDLGDNELPI---DV 53
DB 912 KNRBSSTVSTNSIFYSPLKRGNSKRVYSTSAD-IEENDITRADAPPMDSDD 970
QY 54 OSATOSASTDTANPLDEHEPELYTALENK--TMLNSALNODIMRLACYDTLVHGTP 111
DB 971 SDDSSSSDDIIMSKKTAPE---TNNENKDKESDSSSTHSDEIR---YDS---QYQ 1019
QY 112 AVIKTRSRIRLEDTIMQTIKGRPOVYVOETDPIFLMGNEKGLTKKAKOLEYAKQET 171
DB 1020 DKMERK-----MTREPSPEVYQ-NLEKFFPRALDKRIT-----EGIASPTS 1061
QY 172 PLTSLFDDLRNNT-----PLWSSRPHPMYVLPIMHG-----KPNRSPNTP----- 213
DB 1062 PSLDLSLSPKVVASRTPSPRPVPPDSSEFIQDOLNGKNKRLNDAKPRKRTJR 1121
QY 214 --SHEARQTPNEFRAPELKFOVSVKVAE--DLMGTSDLMEGYTQOSHMOIFGKNSR 270
DB 1122 TTAHEEHLARKN-----SVKLKRONTKMGT---RMVEVENHNVSSINKAKNSK 1167
QY 271 PRVVDYQEIFLTQPVYSDLPMDGKVRMIGMA-----VHNSGESAKLSR----- 317
DB 1168 -----GEYKEFAM-MKGEIMIGSGFAGVYLLCLVTTGEMAAKQVEVPKY 1211
QY 318 -SMNRAYLMAGNEMKNLTVMPRIMGRIFKEGSGSQDPDNDILDYGYGDPVRELYOLENK 376
DB 1212 SSQNEAIIIS-----TV-----EALRSEVSTLKDLHLNITVQYLG-----ENK 1249
QY 377 SWISGTVRNPNRSGKALQDLDVYPLGKISGTFQIFQGYGSLDIYNEHATSFGYGL 434
DB 1250 NNIY-----SLPLEYV--AGGSVGLIRMYGRFDEPLI--KHLTVOVLKGL 1291

```

```

RESULT 7
US-09-463-488A-63
; Sequence 63, Application US/09463488A
; GENERAL INFORMATION:
; APPLICANT: Okamoto, Hiroaki
; APPLICANT: Nishizawa, Tsutomu
; TITLE OF INVENTION: Non-B, Non-C, Non-G Hepatitis Virus Gene, Polynucleotide, Polypeptide, Virus Particle, Method for Isolating Virus Particle, and Method
; FILE REFERENCE: 45455-239360 20111-0040
; CURRENT APPLICATION NUMBER: US/09/463,488A
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: PCT/J998/03340
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: JP 10-82952
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: JP 9-314196
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: JP 9-233246
; PRIOR FILING DATE: 1997-07-25

```

```

; NUMBER OF SEQ ID NOS: 64
; SEQ ID NO: 63
; LENGTH: 761
; TYPE: prt
; ORGANISM: non-B, non-C, non-G hepatitis virus
US-09-463-488A-63

```

```

Query Match          4.0%; Score 94; DB 5; Length 761;
Best Local Similarity 20.7%; Pred. No. 0.77;
Matches 98; Conservative 62; Mismatches 178; Indels 136; Gaps 27;

```

```

QY 30 PVAFVDEVSNDLGDNELPIDVOSATOS-----ASTDTANPLDEHEPELYTALENK 83
DB 229 PKLFEDKWTQODL---CKVPLVITATADLRYPSPCSPTNNPC-----TFQVLRKN 280
QY 84 TMLNSALNODI---MRLACYDTLVH-----GEPRAIK--TKRSIRLDETIQ 128
DB 281 NTVIGTSVAKDQESTQDEEMLYKTDSHYQFATEAOLGRIPAPNPDGTYNT---KQOSWQ 337
QY 129 ---TIKGRPOV---VYQETDPIFLMGNEK--GMLTKKAKOLEYAKQFTPLSLFDL 179
DB 338 DNMSKNSPNTGNSGTYPTTSEMYKIPYDSNNGFPTIRAKQYILERRQ-----CNFN 392
QY 180 DRNNTPLMSSRPHPMYVLPIMHGKPNRSPNTPSHARQF-----TPNEFRAP 229
DB 393 EVNN-----PVSKKVPQPSPTTPTVDYEHGCMFENIFIGPNRY---N 434
QY 230 LKQVVS-----YKVAEEDLW-----GTDSD-----TW---FGYTQ 258
DB 435 LQFQTAIVDTTPPLMDKGNKIMFOYLSKKTQDNEKQCYTLEDMPLMAICFGTYD 494
QY 259 SHQIENGKNSRFRVHDYQEIFLTQPVYSDLPMDGKVRMIGMAVHNSGESAKLSRS 318
DB 495 VERQL--GRVVD---HEHAGLIMICP-YTORPMYDKRPNMGVYVYOTNFGNGMPSG 547
QY 319 WNRAYLMAGNEMKNLTVMPRIM-----GRIEFGSGSQDPD--NDILDYGY-----G 365
DB 548 SCQVPVYWCQRR-----PALMFQOQOVLNDSITGPRAYDEXKYNQVLTLYYNEIFNMWG 602
QY 366 DVAFELYOLEKMSISGTVRNPNRSGKALQDLDVYPLGK---ISGYFQIFQGY 416
DB 603 DMYPVOVKNPCGDSGIV---PQSGRTREVQVYVPLSMGPAYIFHFDSRRGF 653

```

```

RESULT 8
US-09-897-516-7203
; Sequence 7203, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215, 161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO: 7203
; LENGTH: 2590
; TYPE: prt
; ORGANISM: Xenorhabdus sp.
US-09-897-516-7203

```

```

Query Match          3.9%; Score 91; DB 5; Length 2590;
Best Local Similarity 20.9%; Pred. No. 8.7;

```

```

Query Match          3.8%  Score 88.5;  DB 5;  length 742;
Best Local Similarity 24.2%  Pred. No. 2.4;
Matches 86;  Conservative 41;  Mismatches 136;  Indels 93;  Gaps 23;

0Y  47  NELLIPDVQSAQSS-----PTTAPRDLDEHEELITTALENTMLNCALNO---DIMR 98
    ||| : ||| | ||| ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

```

	Query Match	3.8%;	Score 88.5; DB 5; Length 742:
	Best Local Similarity	24.2%;	Pred. No.2.4;
	Matches 86; Conservative		Mismatches 41; Indels 93; Gaps 23.
OY	47 NELPIDVOSATQAS-----TDTANPLDEHBEPLYALENKMTLINC SALNQ----DIMR		98
Dd	187 NEHPARKLPNGGFASHFGYTGRKANPRDDKEG-----LVGMAGLEQTNYDILLS		234
OY	99 ----LACDYTLNHGET-PAVIKTAKSIRLDLETIMOTIGCRFVVUUYQEITTPRLFMGNER-		152

Db 235 GTDGRVYFEKDIYGNALPCTVAEKKAVDGDITYTTLSR-----LQNTLEDLMTQVNEYK 290
Qy 153 -----GMLTKKDAKOLEYAAKQFTPLSLSPD-----LDRNNTPLMSS-----RPHNPMY 196
Db 291 EYVSMATML-NEAKTGEIVAMSSQRP---TFNPETKQGLDDNGT--NQGLLVESPEPES 343
Qy 197 VLPITMHGKPNRSPNTPPSHARQFTPNNEERAPLKFQVSVKKAADLMTGTSDLMEGTYT 256
Db 344 TKLRF-----TTAASMEGQFNPENLEFNRYGSIQVQ-----DVTVNHDH-----YT 384
Qy 257 QGSHMIFNGKNSRPFV-VHDYQPEIFLFL--QPVYSDLPMDGKVRMIGAVHHS--NGE 311
Db 365 R-----LNGKXYLYNRAOIAISSNIGVAKLQKMGDEKMMEXLKFEGFTSTHSLSGE 438
Qy 312 SA-KL--SRSMNRAYLMAGMEKNLTV---MPLWGRIFKEGSGSQPDNDPILD 360
Db 439 SAGKLPCTNFEVDRAMSAFG---QAITVTNFOAMKGFSAIANGSMLOPHYISKIYD 491

RESULT 11
US-09-897-516-7356
Sequence 7356, Application US/09897516

GENERAL INFORMATION:

APPLICANT: Corbin, David R.
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Huesling, Joseph E.
APPLICANT: Krasomil-Osterfeld, Karina C.
APPLICANT: Malvar, Thomas M.
APPLICANT: Slater, Steven C.
APPLICANT: Spilidonov, Sergei
TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
FILE REFERENCE: 38-21(51847)B
CURRENT APPLICATION NUMBER: US/09/897,516
CURRENT FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/215, 161
PRIOR FILING DATE: 2000-06-30
NUMBER OF SEQ ID NOS: 8409
SEQ ID NO 7356
LENGTH: 678
TYPE: PRF
ORGANISM: Xenorhabdus sp.
US-09-897-516-7356

Query Match 3.7%; Score 86.5; DB 5; Length 678;
Best Local Similarity 20.0%; Pred. No. 3.2;
Matches 56; Conservative 32; Mismatches 85; Indels 107; Gaps 13;

Qy 128 QTIKGRVYVQETP--DPIF--LMGNKGM-----TKKDAKOLEYAAKQFTPL 173
Db 62 KTIPIGFSQVRSCTGMDPVRFGMFGSRLMLDDGEVMSGSCSRMDAPSSYISPESYDVL 121
Qy 174 SLSPDLRNNPFLMSSRPINPMYVLPITFMHGKPNRSPNTPPSHARQFTPNNEERAPLKFQ 223
Db 122 SL---VKGFQTVLMG-----PNSAGTIRERSR---PQSEPAKNA 158
Qy 234 VSVKKAADLMTGTSDLMEGTYQOSHMOIFNGKNSRPFVHDYQPEIFLQVYSDLP 293
Db 159 IST-VAGSGNRDGNADLSFG--NQYGYMLIGKNS---RSNDYK----- 197
Qy 294 DQKVRMIGAVHHSNGESAKLSRSNMRAYLMAGMEKNLTVMPRIWGRIFKEGSGSQPD 353
Db 198 -----DGDGYRIPSHMNK-----WNTDVAL-----GWTDP 222
Qy 354 DNPDLIDYVYGDVRFVLELKNKSNISGTVRYNPRSGKA 393
Db 223 DNTLLELANGGD-----GKARYAGRAMDS 248

RESULT 12
PCT-US01-27760-801
Sequence 801, Application PC/TUS0127760

GENERAL INFORMATION:
APPLICANT: Hysq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-115/ 785
CURRENT APPLICATION NUMBER: PCT/US01/27760
CURRENT FILING DATE: 2001-10-11
PRIOR APPLICATION NUMBER: 09/687,527
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 892
SOFTWARE: Custom
SEQ ID NO 801
LENGTH: 1021
TYPE: PRF
ORGANISM: Homo sapiens
PCT-US01-27760-801

Query Match 3.6%; Score 84.5; DB 1; Length 1021;
Best Local Similarity 19.7%; Pred. No. 8.8;
Matches 68; Conservative 44; Mismatches 100; Indels 133; Gaps 17;

Qy 131 KGRVYVQETTPDPIFLMGNEKGMILTKDAKOLEYAAKQFTPLSLSPDLDNRNTPMW-SS 189
Db 499 KAKAQVYLQRLD-----VQEIFHKROYSLMKLAKQTRPVQ--PVAPRPSSPKVSS 550
Qy 190 RPHNPMYVLPITFMHGKPNRSPNTPPSHARQFTPNNEERAA---PELKFQVSVKKAADL 244
Db 551 KTSQPSSTVPL--ARPLRTSEBP-----YFTELNSRGEDEDEKFEVK-----SEEL 596
Qy 245 WGTSDLMGTYQOSHMOIFNGKNSRPFVHDYQ-----ELFLQPVYSDLPMDGKVR 298
Db 597 F-----ESHHERNPPELEQOARIGDISPRRRIITLLETEELY----- 634
Qy 299 MIGAVHHSNGESAKLSRSNMRAYLMAGMEKNLTVMPRIWGRIFKEGSGSQPDNDPDI 358
Db 635 -----IKEIKIIDGY-----ITPMEDIW---LKHILPDLQNNKDF 668
Qy 359 LDYVYGDVRFVLELQLEN-----KSNISGTVRYN---PRS-- 389
Db 669 L---FGNIRELYEFHNRFTLKELEKCAENPELLAHCFLEKRDLOIYFYKHRLPRARA 724
Qy 390 -----GKGAQLDQVYVPLGKGISGTFQIFGQYGSLLDY 423
Db 725 IMQECDCAYEGVQOROLDHNLPL-----FYLKGPSQRLIKY 762

RESULT 13
US-09-897-516-6909
Sequence 6909, Application US/09897516

GENERAL INFORMATION:

APPLICANT: Corbin, David R.
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Huesling, Joseph E.
APPLICANT: Krasomil-Osterfeld, Karina C.
APPLICANT: Malvar, Thomas M.
APPLICANT: Slater, Steven C.
APPLICANT: Spilidonov, Sergei
TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
FILE REFERENCE: 38-21(51847)B
CURRENT APPLICATION NUMBER: US/09/897,516
CURRENT FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/215, 161
PRIOR FILING DATE: 2000-06-30
NUMBER OF SEQ ID NOS: 8409
SEQ ID NO 6909
LENGTH: 802
TYPE: PRF
ORGANISM: Xenorhabdus sp.
US-09-897-516-6909

Query Match 3.5%; Score 83.5; DB 5; Length 802;

Best Local Similarity 19.0%; Pred. No. 7.7;
Matches 67; Conservative 59; Mismatches 109; Indels 117; Gaps 20;

```
QY 147 LMGNEKMLTKKADAKOLEYAKOFTPLSLFDDRRNTPLMSSRPHNPMYLPFMHGRP 206
      ||| : : : : : ||| : : : : : ||| : : : : :
Db 41 LVGLEKALVSEHNPVYKAEK----MGKLFDDIKINNLTNT---EDIALANFL----- 89
QY 207 NR-----SPNTSHEAROTPNDEFRAPELKFQVSVKKAEDLMGTSLSMGSTYQOSH 260
      ||| : : : : : ||| : : : : : ||| : : : : :
Db 90 TRLLCFEADTGIIFSDAQTS-----AIKSYTEED--GSDLLAFL----- 128
QY 261 WOLFNGKNSRPFVHQPEIFLTPQPYSDLPMDGKVRMIGMAVHNSGESAKLSRWN 320
      ||| : : : : : ||| : : : : : ||| : : : : :
Db 129 YQFVSVALNT-----PESDTINLPKHL-----NNFRYN-GDLFSSDDPIPELKGKR 176
QY 321 RAYLMAG-MEMKNLYTPRMGRIFKSGSGSDPDNDPILD-----YVGGDVRF 369
      ||| : : : : : ||| : : : : : ||| : : : : :
Db 177 RILLDCGSMDSMSEIN--PDIFGSMFQA-----VIDEGRSLGQHYTSYNIIMK 223
QY 370 LYQ-----LEKNSISGTVR-----YNPSSGKG-----ALQLD 397
      ||| : : : : : ||| : : : : : ||| : : : : :
Db 224 VIQPLEPLRSDPEKRRKRVNGLAKLVRIKIKYDFPACGSGCNFLIVAYKELRLKLEME 283
QY 398 YVYPLCK-----GISGYFOIFQGYGOSLIDYHNHEATSFVGL-----MLNDM 439
      ||| : : : : : ||| : : : : : ||| : : : : :
Db 284 IYALGEIDPQSIIPMSG-LHLSQFYGIEIDDFACEVARLSLMLAEHQJLNNQW 334
```

RESULT 14

US-09-897-516-5307

Sequence 5307, Application US/09897516

GENERAL INFORMATION:

APPLICANT: Corbin, David R.
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Huesling, Joseph E.
APPLICANT: Krasomil-Osterfeld, Karina C.
APPLICANT: Malvar, Thomas M.
APPLICANT: Slater, Steven C.
APPLICANT: Spiridonov, Sergei
TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
FILE REFERENCE: 38-21(51847)B
CURRENT APPLICATION NUMBER: US/09/897,516
CURRENT FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/215, 161
PRIOR FILING DATE: 2000-06-30
NUMBER OF SEQ ID NOS: 8409
SEQ ID NO 5307
LENGTH: 595
TYPE: PRF
ORGANISM: Xenorhabdus sp.
US-09-897-516-5307

Query Match 3.5%; Score 82.5; DB 5; Length 595;

Best Local Similarity 23.4%; Pred. No. 6.1;

Matches 50; Conservative 29; Mismatches 76; Indels 59; Gaps 13;

```
QY 30 PVAFFVEVRSENDLGONDELPIDVOSA---TOSASTDTANPLDDEHPELYTTALENKT 85
      ||| : : : : : ||| : : : : : ||| : : : : :
Db 186 PSSAVNTSTVDT-----SVDTNSVDSITTAIVETDQANTADTTETDM-----TGQF 232
QY 86 LINCASLADIMRLACYDT-----LVHGETPAVITKTRISIRLDEITWQING---KP 134
      ||| : : : : : ||| : : : : : ||| : : : : :
Db 233 TINSVAPNHQVVKI--YDTQSOPLANLLAQAEQGV-----TLVVGPLLR 276
QY 135 QV--VQOETDPTFLMGNEKMLTKKADAKOLEYAKOFTPLSLFDDRRNTP--LMSSR 190
      ||| : : : : : ||| : : : : : ||| : : : : :
Db 277 QVEQLVOINT-PLNTLA-----LNKPDPTQLH--PNFCYFSLSPEDAKSMAHQIWOQ 327
QY 191 PHNPMYVLPFMHGRKPSRSPNTPSHEAROPTPNE 224
      ||| : : : : : ||| : : : : : ||| : : : : :
Db 328 KHNPLVIVPRSVLG--SRVANAFATEWQJLGSGE 359
```

RESULT 15

US-09-800-198-72

Sequence 72, Application US/09800198

GENERAL INFORMATION:

APPLICANT: Vernet, Cornie AM
APPLICANT: Fernandes, Elma
APPLICANT: Shimkels, Richard A
APPLICANT: Hermann, John L
APPLICANT: Majumder, Kunud
APPLICANT: Mishra, Vishna
APPLICANT: Mezes, Peter S
APPLICANT: Rastelli, Luca
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 15966-697
CURRENT APPLICATION NUMBER: US/09/800,198
CURRENT FILING DATE: 2001-03-05
PRIOR APPLICATION NUMBER: 60/186,596
PRIOR FILING DATE: 2000-03-03
NUMBER OF SEQ ID NOS: 98
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 72
LENGTH: 2765
TYPE: PRF
ORGANISM: Rattus norvegicus
US-09-800-198-72

Query Match 3.5%; Score 82.5; DB 5; Length 2765;

Best Local Similarity 20.4%; Pred. No. 58;

Matches 99; Conservative 58; Mismatches 156; Indels 173; Gaps 27;

```
QY 57 TOSASTDTANPLDEH-----EPELYTTALENKTMLINCASLADODI 96
      ||| : : : : : ||| : : : : : ||| : : : : :
Db 1609 TYSADNDVTELIDNNNGSLKIRDSGCMRHLMDNDQIITLVGTNGLAIVSYQNE- 1667
QY 97 MRLACYDPTLVHGETPAVITKTRISIRLDEITWQI-----KKPOVYQET 141
      ||| : : : : : ||| : : : : : ||| : : : : :
Db 1668 LGLMTYD---GNT-GLIATKS---DEGTWTFYDYDHEGRLLNVTPTGVVSLHREM 1718
QY 142 TDPIFL-----MGNEKMLTKKADAKOLEYAKOFTPLSLFDDRRNTPLMSSRPHNPM 195
      ||| : : : : : ||| : : : : : ||| : : : : :
Db 1719 EKSTIVDIENSNRDNDVITNLSSEASTYVQ-----DOYRNSYOLCSNKTLYM 1770
QY 196 YV--LPIFMHGRKPSRSPNTPSHEAROPTPNDEFRAPELKFQVSVKKAEDLMGTS 253
      ||| : : : : : ||| : : : : : ||| : : : : :
Db 1771 YANGMGVSEHSEPHYLAQT-----LPTIGRC-----NLSLPMEN-----GLNSIEMR 1813
QY 254 GYTOQSHWQIFNGKNS---RPRFVN-----DIQPELFLTPQVYS-----LPMGK 296
      ||| : : : : : ||| : : : : : ||| : : : : :
Db 1814 LRKEQ-----IKGKYTLIGRRLRVHGRNLLSIDYDRI-RTREKIYDDHRKFTLRIT 1867
QY 297 VRMI-----GMAVNHNS---NGESAKLSRSMNRRAYLMAAGEMKNLYTPMIRMGRIFK 347
      ||| : : : : : ||| : : : : : ||| : : : : :
Db 1868 GRPLMLPSSGLAAVWVSFFNGRLAGLQRC-----AMSEKRDIDKQGNIVSRMPADG 1920
QY 348 SGSDPDNDPILDYGYGDV-----RFLYOLEKNSNIS----- 380
      ||| : : : : : ||| : : : : : ||| : : : : :
Db 1921 K-----VMSYSLYDKSMVLLLOQROXYIFEDYSSDRLAHVMTMPVARHSMSTHTS 1970
QY 381 -GTVR--YNPSSGKALOLDY-----VYPLKGISGYQIIFQGYQO---SLIDYNH 425
      ||| : : : : : ||| : : : : : ||| : : : : :
Db 1971 IGYIRNIYNPESNASVIFDYSDGDIRLKTSLFTGT---RQVFKYKGLSKLSLSEIYVDS 2026
QY 426 EATSG 431
      ||| : : : : : ||| : : : : : ||| : : : : :
Db 2027 TAVTFG 2032
```

Search completed: November 30, 2001, 14:25:08
Job time: 540 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:26:27 ; Search time 77.9 Seconds
(without alignments)
829.941 Million cell updates/sec

Title: US-09-787-083-6
Perfect score: 2360
Sequence: 1 MKVSLSTLTLSILSCFALIA.....YNHEATSPGCLMLNDMGL 442

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP:REMBL.17:*
2: SP:archaea:*
3: SP:bacteria:*
4: SP:fungi:*
5: SP:human:*
6: SP:invertebrate:*
7: SP:mammal:*
8: SP:organelle:*
9: SP:phage:*
10: SP:plant:*
11: SP:rodent:*
12: SP:virus:*
13: SP:vertebrate:*
14: SP:unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	815	34.5	382	2 09K0U7	09K0U7 neisseria m
2	815	34.5	409	2 09J721	09J721 neisseria m
3	373.5	15.8	306	2 09CL22	09CL22 pasteurella
4	366	15.5	329	2 09PMU8	09PMU8 campylobact
5	360.5	15.3	292	2 09SID7	09SID7 yersinia ps
6	360	15.3	289	2 09L6N9	09L6N9 salmonella
7	356.5	15.1	292	2 09Z4N8	09Z4N8 enterobacte
8	342.5	14.5	297	2 03Z349	03Z349 campylobact
9	246.5	10.4	355	2 025241	025241 helicobacte
10	240.5	10.2	355	2 09ZLX5	09ZLX5 helicobacte
11	123	5.2	278	2 09XB53	09XB53 erwina car
12	114	4.8	1686	4 000443	000443 homo sapien
13	110.5	4.7	602	11 063485	063485 rattus norv
14	109.5	4.6	821	2 059241	059241 bacillus sp
15	108	4.6	824	2 09F216	09F216 bacillus sp
16	105.5	4.5	901	5 018749	018749 caenorhabdi
17	104	4.4	783	2 045554	045554 bacillus sp
18	104	4.4	798	4 09UN32	09UN32 homo sapien
19	104	4.4	798	4 09UBK2	09UBK2 homo sapien

20	103.5	4.4	3247	12 065553	065553 bovine herp
21	103	4.4	435	2 09X9C0	09X9C0 streptococc
22	101.5	4.3	403	5 091736	091736 hydra magni
23	101.5	4.3	1046	2 084941	084941 streptococc
24	101	4.3	4307	5 019319	019319 caenorhabdi
25	100.5	4.3	1509	11 061194	061194 mus musculi
26	100	4.2	719	2 057114	057114 streptococc
27	100	4.2	719	2 09RET8	09RET8 streptococc
28	99.5	4.2	719	5 09V9E6	09V9E6 drosophila
29	99.5	4.2	1658	11 061182	061182 mus musculi
30	99	4.2	719	2 054948	054948 streptococc
31	99	4.2	719	2 054946	054946 streptococc
32	99	4.2	719	2 09WVW0	09WVW0 streptococc
33	99	4.2	719	2 09WM11	09WM11 streptococc
34	99	4.2	719	2 09RET6	09RET6 streptococc
35	99	4.2	719	2 09RET4	09RET4 streptococc
36	99	4.2	727	2 09F2G4	09F2G4 streptococc
37	99	4.2	1372	10 09FL92	09FL92 arabidopsis
38	99	4.2	5005	2 09PP25	09PP25 ureaplasma
39	98.5	4.2	467	4 09UMT5	09UMT5 homo sapien
40	98.5	4.2	467	4 09BYT2	09BYT2 homo sapien
41	98.5	4.2	576	2 09L115	09L115 streptomyce
42	98.5	4.2	660	2 09KKG8	09KKG8 bacillus ha
43	98.5	4.2	788	10 09SCV5	09SCV5 arabidopsis
44	98	4.2	719	2 054947	054947 streptococc
45	98	4.2	719	2 054949	054949 streptococc

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	382 AA.
09K0U7	09K0U7			
ID	09K0U7			
AC	09K0U7			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	PHOSPHOLIPASE A1, PUTATIVE.			
GN	NMB0464.			
OS	Neisseria meningitidis (serogroup B).			
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.			
OX	NCBI_TaxID=491;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-MC58 / SEROGROUP B;			
RX	MEDLINE=2017575; PubMed=10710307;			
RA	Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J., Nelson W.C., Gwinn M.L., Deboy R., Peterson J.D., Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A., Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B., Cotton M.D., Uitterlinden T.R., Khouri H., Qin H., Vamathevan J., Gill J., Scarlato V., Maignani V., Pizsa M., Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappelli R., Venter J.C.; "Complete genome sequence of Neisseria meningitidis serogroup B strain MC58."			
RT	Science 287:1809-1815(2000).			
RL	EMBL: AE002403; AAF40901.1; -.			
DR	TIGR: NMB0464; -.			
DR	InterPro: IPR003187; PLAI.			
DR	Pfam: PF02253; PLAI: 1.			
KW	Complete proteome.			
SO	SEQUENCE 382 AA; 42714 MW; B468A802F062E836 CRC64;			

Query Match	34.5%	Score 815;	DB 2;	Length 382;
Best Local Similarity	44.7%	Pred. No. 1.6e-62;		
Matches 163;	Conservative 67;	Mismatches 111;	Indels 24;	Gaps 8;
Oy	87	INCAALNDIMRLACVDTLVGERTPAVI-----TKRSIRLDETITWOTI-KGKPOVYVOE 140		
Db	33	LOCALINDVTRLACVDRIFRAQLPSSAGDEGSKAVNLTEIVRSLSKGEAVIYVEK 92		

OY	141	TTDPFLMGNGENKGMTLTKDKDALEAAQOFPLTSLSFDLRNN-SPMSSRPHPNMYLP	199
Db	93	GGDAL-----+-----PADSAGETADITPPLSLMTYDLKDKNDLGGILGVRHNPMYLP	139
OY	200	IFMHCKPMSRPTPSHEAR-OETPNEFRAPELKFOVSVYKAAEDLMGTSDLMFGYTOQ	258
Db	140	LMYNSSPNVAPGSPRGTTVDEKFGQOKRAETKLGQVSEFSKLAEDLEFKTRADLMFGYTOR	199
OY	259	SHMOLF-N-GKNSRPFRRVNDYQPEIFILTPQVYDDLPMWDCVNRITIGMAVHNSGESAKLSR	317
Db	200	SDMOQLYNGRKAAPRNMDYKKEIFILTPQVKADLPFGGRRLMGLAGFVHQSGQSRPESR	259
OY	318	SNRRAYLMAAGMEWKNULTWMPRIINGRIKEFGSSQDDNDPIDYGYGVRFLYOLENS	377
Db	260	SNNRIRYMAAGMGKGLTYIPRYVVAFAFD-SSDK-NDNPDIADYMGYDQKLTQYRLNDRQ	317
OY	378	NISGTVRYNPSRSGKALDLDVYPLGKGISGYFQIFQGGSLIDYNHETSFGVGLMLN	437
Db	318	NYSVLRVLPKGYGAIEAAVYFPIKGLKGVRRGFHGESLIDYNHKONGIGIGLMPN	377
OY	438	DWNGL 442	
Db	378	DLDGI 382	
RESULT	2		
ID	09JUT21	PRELIMINARY:	PRT: 409 AA.
AC	09JUT21.		
DT	01-OCT-2000 (TREMBLrel. 15, Created)		
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	PURATIVE PHOPHOLIPASE.		
GN	NMA2021.		
OS	Neisseria meningitidis (serogroup A)		
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.		
OX	NCBI_TaxID=65699;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=22491 / SEROGROUP A / SEROTYPE 4A:		
RX	MEDLINE=20222556; PubMed=10761919;		
RA	Parikh11 J., Achman M., James K.D., Bentley S.D., Churcher C.,		
RA	Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,		
RA	Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,		
RA	Jegels K., Leather S., Moule S., Mungall K., Quail M.A.,		
RA	Whitehead M.A., Rutherford K.M., Simmonds M., Skelton J.,		
RA	Whithead S., Spratt B.G., Barrett B.G.		
RT	"Complete DNA sequence of a serogroup A strain of Neisseria		
RT	meningitidis 22491."		
RL	Nature 404:502-506(2000).		
DR	EMBL: AL162757; CAB85240.1; -		
DR	InterPro: IPR003187; Pfam: PF02253; Pfam: PF02253; Pfam: PF02253.		
KW	Complete proteome.		
SO	SEQUENCE 409 AA; 45862 MW; CD6585B064D01A41 CRC64;		

	Query Match	34.5%	Score 815;	DB 2:	Length 409;
	Best Local Similarity	44.7%;	Pred No. 1.8e-62;		
	Matches 163; Conservative	67;	Mismatches 111;	Indels 24;	Gaps 8;
Oy	INCSALNODIMLACVDYLHGEPFAVI-----TKRSIRLDLETIMORI-KGKPOVYOE	140	:	:	:
Dd	LQCALTLTDNVLTACTDRIFAQLSSAGQEQEGESAVNLTEFYRSSIDKEAYIVKER	119	:	:	:
Oy	141 TTDFIFLMGNCKGLTKKDAAKOLEYAANKOTPLSLSEFDIDRNN-TPLWMSRPHPNYVL	199	:	:	:
Dd	120 GGDAL-----PADSAGEATADITYTLPISLVYDLDKKDLRGILLGVREHNPYMLP	166	:	:	:
Oy	200 IHHMKRNPSPTSPHSHEAR-QETPNDEFRRPELKFOYSYVKVAEDLDMGSDSLMFGYTQQ	258	:	:	:
Dd	167 LMYNNSPNAPASPRGGTTVOEKFSQQAERFKLOVSFESKLAIEDLFKTRADLMVGCTYOR	226	:	:	:

[illegible]

Query Match	15. 8%	Score 373.5;	DB 2;	Length 306;
Best Local Similarity	37.6%;	Pred. No. 2,4e-24;		
Matches 86;	Conservative 43;	Mismatches 91;	Indels 9;	Gaps 5;
QY 214	SHEAQGFPPNE-FRAPELKFQVSVKKAAMEDLMDGTDSDLMFGYTOOSHQIFENKKSRRPF	272		
DB 81	TVSNRHFELKRETKQXDEIKFISTLAPLMRGILGNNSVLAASITQKSWQLSNVDDSSPF	140		
QY 273	RVHDYQPEIFELQAPVYSDLPMDGKVRMLICGAVHHNSNG--ESAKLSRSWNRRAYLAMGEM	330		
DB 141	RETNREPOLFLAMKQYSLRPFQWTLQDVETGLNHOSNGDDAEKLSRSMRLLYVASAIK	200		
QY 331	KNLTYMPRIWGRIFKEGSGSQDPDNDPDLIDYGYGDVRF-LYQLENKSNISETVRNPRS	389		
DB 201	QMWTEYIKRPMWRIPEK---AKNDNDPDIKKYGHGFHDVALGYYYHGHQFLSG--HYNPIS	255		
QY 390	GKGALQLDVYVPLGKIGISGFQIFQGYGOSLIDYHNHATSFEGVGLMND	438		
DB 256	NKGGLAASYSPYITKKNIRFQYNGYGSGLIDYQOORIGRIGISILNN	304		
RESULT 4				
Q9PMU8				
ID Q9PMU8	PRELIMINARY;	PRT;	329 AA.	
AC Q9PMU8;				
DT 01-OCT-2000 (TREMBLrel. 15, Created)				
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)				
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)				

QY	345	KEGSSQSDDDNDIDYGYGDFRFLYQLENSNJSQYRYPBRSKGALQDIDYYPICK	404
Db	200	-----GSDNDNDIDIKKMGYQYQKLGYYH-GGAVLSAAGQYNNWNTGYGGAEEVGLSPYTK	253
QY	405	GISGYEQIFQGYGQSLIDYNHEATSFYGVGLMND	438
Db	254	HVRLTYQYSGESGLIDYNFMQTRVGVGMND	287
RESULT	7		
Q9Z4N8	PRELIMINARY:	PRT:	292 AA.
AC	Q9Z4N8:		
DT	01-MAY-1999 (TREMBLrel, 10, Created)		
DT	01-MAY-1999 (TREMBLrel, 10, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel, 17, Last annotation update)		
DE	OUTER MEMBRANE PHOSPHOLIPASE A PRECURSOR (Ec 3.1.1.32).		
GN	PLDA.		
OS	Enterobacter agglomerans.		
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		
OC	Pantoea.		
OX	NCBI_TaxID=549;		
RN	(1)		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=94131966; PubMed=8300539;		
RX	Brok R.G., Brinkman E., van Bostel R., Bekkers A.C., Verheij H.M.,		
RA	Tomassen J.;		
RT	"Molecular characterization of enterobacterial plda genes encoding		
RT	outer membrane phospholipase A.;"		
RL	J. Bacteriol. 176:861-870(1994).		
RN	(2)		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=99120145; PubMed=99251577;		
RX	Brok R.G., Boots A.P., Dekker N., Verheij H.M., Tomassen J.;		
RA	"Sequence comparison of outer membrane phospholipases A: implications		
RT	for structure and for the catalytic mechanism.;"		
RL	Res. Microbiol. 149:703-710(1998).		
DR	EMBL; AF034414; AAD03498.1; -.		
DR	InterPro: IPR003187; PLAL.		
PF	Pfam: PF02253; PLAL1.1.		
DR	PRINTS; PR01486; PPHPLIPASEAL.		
KW	Signal; Hydrolase.		
FT	SIGNAL	1	20
FT	CHAIN	21	292
SO	SEQUENCE	292 AA; 33719 MW; B75516D093B2BEEA CRC64;	
			POTENTIAL.
			OUTER MEMBRANE PHOSPHOLIPASE A.
	Query Match	15.1%;	Score 356.5; DB 2; Length 292;
	Best Local Similarity	40.1%;	Pred. No. 6,8e-23;
	Matches	85; Conservative	31; Mismatches 87; Indels 9; Gaps 5
QY	229	ELKQVSVYKRAEDLMGDSIDLMFCYQQQSHWQIFNGKNSRPFRRHDDQPELEFLQPY-	287
Db	86	EKFQFLSLGEPIMKCIAGNSLGLGASYQSRWQASNSDESPFRTEYNPEQIFLAWARD	145
QY	288	YSDLPWDGKVRNIGGAVHNSGESAKTISRNNRARIYLAGMGMKNTVPRIRMGRTFKEG	347
Db	146	YELAWG--TFREVEEGYNNQNSGKADPTSRSDRYTRMLAQRGNLEIDLKPYRIRPESD	203
QY	348	SGSQPDNDPIDIDYGYGDFRFLYQD-ENKSNISGTVRYRNPBSGQALQDLYVPIGKCI	406
Db	204	S---KDDNPDIKKMGYQYQKLGYYH-GGAVLSAAGQYNNWNTGYGGAEEVGLSPYTK	258
QY	407	SGYEQIFQGYGQSLIDYNHEATSFYGVGLMND	438
Db	259	RYTYQVSGYSGESMIDYRQTRVGVGMND	290
RESULT	8		
Q3Z349	PRELIMINARY:	PRT:	297 AA.
AC	Q3Z349:		
AC	032349;		

Query Match	14.5%	Score 342.5	DB 2	Length 297
Best Local Similarity	32.4%	Pred. No. 1.1e-21		
Matches 97	Conservative 37	Mismatches 108	Indels 57	Gaps 11
OY	178	DLDRNNTPLMSRPNMXYLPIFMHGKPNRSPNT	-----P	213
DB	16	DLKENNASLSLRKHETQ	-----NTQKPTSTKEDFSRLALNLYGENSSFNPLGIS	66
OY	214	SHEARQFTP	-----NEFRAPELKFOVSKVAAEDLMGTDSDLMFGYTOQSHWQTF	264
DB	67	SYKMYVFLPFAVSFSGSLGGENKTEMKRQLSIKKHLFDLLGLGKYYGYGTSMWQ	--	124
OY	265	NGKNSRPRVADYOPEIFLTQPV	-YSDLPMQCKVMRMIGMAVHSHNGESAK	--LSRSNKR 321
DB	125	NKXHSPPFETNYOPEFFDIPDLFHEEDYKFLNNLR	---VGLIHSNGKDEMLERSNMR	181
OY	322	AYLMGMEKNNLTWPRIWRIFFKGGSGSOPDNPDILDYGYGVDRFLYOLEKNSNIG	S	381
DB	182	IYASSVFLKQRLPEPRIMYRI	---PENSEDDNNEPIHTYMKNFQIN	-IGSLGNDYFIUL 237
OY	382	TVRYNP	-RSGKALQLDLYVPL	-GKGISGYFOIFOGYGOSLIDYNHEATSEFGVGLMLN 437
DB	238	MLRNMLDFHNDKGAQVVDIGDYDIFDNGIYMYLOYENGYSGLIDYNNKRLQRLSTAFLLS		296
RESULT	9			
O25241		PRELIMINARY:	PRT:	355 AA.
ID	O25241			
AC	01-JAN-1998	(TREMBLrel. 05, Created)		
DT	01-JAN-1998	(TREMBLrel. 05, Last sequence update)		
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)		
DE	PHOSPHOLIPASE A1 PRECURSOR (DR-PHOSPHOLIPASE A).			
GN	HP0499.			
OS	Helicobacter pylori (Campylobacter pylori).			
OC	Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;			
CC	Helicobacter.			
OX	NCBI_TaxID=210;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-26695 / ATCC 700392;			
RX	MEDLINE-97394467; PubMed-9252185;			
RA	Tombl J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,			
RA	Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,			
RA	Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,			
RA	Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,			
RA	McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.R.,			
RA	Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,			

RA Cotton M.D., Meldman J.M., Fujii C., Bowman C. Matthew L., Mallin E.
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Velter U.C.;
RT "The complete genome sequence of the gastric pathogen *Helicobacter*
RT *pylori*."
RL Nature 368:539-547(1997).
DR EMBL: AE000564; AAD07564.1; -.
DR TIGR: HP0499; -.
DR InterPro: IPR003187; Pfam.
DR Pfam: PF02253; Pfam: 1.
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 355 AA; 42486 MW; 461960F1565AB0A CRC64;

Query Match	10.48;	Score 246.5;	DB 2;	Length 355;
Best Local Similarly	25.18;	Pred. No. 3.2e-13;		
Matches 82; Conservative	41;	Mismatches 117;	Indels 87;	Gaps 11;

Qy	157	KKDAKOLEYAKORTPLSDDDDRNNTPLMSSPHNPMVLYPLRMGKRNBSRNPISHE	216
Db	69	KKYLMNDYLOTLYLPRYHS-----TPEOMYHNIN-----	1020
Qy	217	ARQETPNEFRABELKFOVSVKKAEDLWGTDSDLMFGYTOOSHMOJTEFNKNSRPERVHD	276
Db	103	---YORNEF-----KFOISFVRPVEFRHILMTKGTLYLATYLOTDFQYLYNDPOQAPMRMN	154
Qy	277	KQRELFELTOPYSDLPMDGK---RMJGMGVHNHNSG-ESAKLSRSNRRATYLAAGMNMKN	332
Db	155	FMPELITYPL-NKRPQGGKIGNFSELMIGOMHISNGVGAGOCYOPNK---EGNDENO	2099
Qy	333	LTVMPRI-----WGRIFKEGSGSOP-----DDNDP	357
Db	210	FPQGPVLYIKDYNCGKDYVRWMCGRSVSAGORPVEFLVWEKGLKIMVAYMPVPYDOSNPN	269
Qy	358	ILDYVGGVDFEFL-----OLENKNISGTYVYNNRSKGALOLDYVYPLKGISGYE	410
Db	270	LIDWGYGNNAKIDYRGRGHHHELDYDLFYQWRYD---RHHGAFRLGYTYRINPFGVYA	327
Qy	411	QIFQYGOSLIDYNNHEATSPGVGLMLN	437
Db	328	QMFGNYGDGLYEXDYVFSNIRGTVGIRLN	354

RESULT	10			
Q9ZLX5				
Q9ZLX5				
ID	Q9ZLX5	PRELIMINARY;	PRT;	355 AA.
AC	Q9ZLX5;			
DT	01-MAY-1999	(TEMBLrel. 10, Created)		
DT	01-MAY-1999	(TEMBLrel. 10, last sequence update)		
DT	01-JUN-2001	(TEMBLrel. 17, last annotation update)		
DE	PUTATIVE PHOSPHOLIPASE A1.			
GN	PLA4 OR JHP0451.			
OS	Helicobacter pylori J99 (Campylobacter pylori J99).			
OC	Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;			
OC	Helicobacter.			
OX	NCBI_TaxID=85963;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99120557; PubMed=9923682;			
RA	Aim R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Dolg P.C.,			
RA	Smith D.R., Noonan B., Guild B.C., deLonge B.L., Carmel G.,			
RA	Tumilimo P.J., Caruso A., Ulla-Nickelsen M., Mills D.M., Ives C.,			
RA	Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.			
RA	Trust T.J.;			
RT	"Genomic sequence comparison of two unrelated isolates of the human			
RT	gastric pathogen Helicobacter pylori.";			
RL	Nature 387:176-180(1999).			
DR	EMBL; AE001479; AAD06029.1; -.			
DR	InterPro: IPR003187; PLA1.			
DR	Pfam: PF02253; PLA1; 1.			
KW	Complete proteome.			
Q0	SEQUENCE	355 AA;	42329 MW;	B3CE9810EBA2FDC9 CRC64;

Query Match	10.28;	Score 240.5;	DB 2;	Length 355;
Best Local Similarity	25.18;	Pred. No. 1e-12;		
Matches	83;	Conservative 44;	Mismatches 109;	Indels 95;
				Gaps 13

[illegible]

RESULT	11		
Q9XB53			
ID	Q9XB53	PRELIMINARY:	PRF: 278 AA.
AC	Q9XB53;		
DT	01-NOV-1999 (TrEMBLrel. 12, Created)		
DT	01-NOV-1999 (TrEMBLrel. 12, Last sequence update)		
DT	01-NOV-1999 (TrEMBLrel. 12, Last annotation update)		
DE	KDUI.		
GN	Erwinia carotovora.		
OS	Erwinia carotovora.		
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		
OC	Pectobacterium.		
OX	NCBI_TaxID=554;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-ATCC 39048, GS101.		
RX	MEDLINE=98065591; PubMed=9402024;		
RA	McGowan S.J., Sebahia M., O'Leary S., Hardie K.R., Williams P.,		
RA	Stewart G.S., Bycroft B.W., Salmund G.P.;		
RT	"Analysis of the carbenem gene cluster of Erwinia carotovora:		
RT	definition of the antibiotic biosynthetic genes and evidence for a		
RT	novel beta-lactam resistance mechanism.";		
RL	Mol. Microbiol. 26:545-556(1997).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-ATCC 39048, GS101.		
RX	MEDLINE=98276484; PubMed=9614345;		
RA	McGowan S.J., Bycroft B.W., Salmund G.P.;		
RT	"Bacterial production of carbapenems and clavams: evolution of beta		
RT	lactam antibiotic pathways.";		
RL	Trends Microbiol. 6:203-208(1998).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-ATCC 39048, GS101.		
RA	McGowan S.J.;		
RA	Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.		
EMBL	U17224; AAD38237.1; -;		
SO	SEQUENCE 278 AA: 31341 MW; 63769P4F3550E1B7 CRC64;		

Query Match	5.2%;	Score 123;	DB 2;	Length 278;
Best Local Similarity	21.6%;	Pred. No. 0.012;		
Matches 74;	Conservative 52;	Mismatches 128;	Indels 88;	Gaps 16

OY	51	IDVOSATQSTADTPANLDEPELYTAAENKMLINCALNDIMRIACYPDLVGET	110
Db	1	MDVQSVHS-----EHAKLDTLTLEKKFLIQDITFPNOYTWTSHIDRIVG--	48
OY	111	PAVIKTNRKSIRLDFTIMO-----TIKGDPVVYQETDPIFLMGNEKM	154
Db	49	-GIMPVDGEITFDGIGQGVNVELERRBELGLINIGSPAKVIDGS---YEYGNEL	104
OY	155	LTKDAQOLEFAAAGFPLSLISFDLRNNPFLMSSRRNPAYVLPTMHGKPNRSPTPS	214
Db	105	YVGGAKRLAS-----SLDSAPKALYLNSAPAHAVFPRITODATAIAPGDV	155
OY	215	HEAROFTENERAPELKRQVSVKRAKAEDDLWGTDSDLMFGYT---OOSHMOIFNGKSRP	271
Db	156	KTCNKRTICKLIVEV-----VETCO-----LSMGLTLRAESSNW-----NSMP	194
OY	272	FRVHDYOPEIFLTGPVYSDLPEWDKVRMIGMAVHNS----NGESAKLSKNWRAYLMA	356
Db	195	THTHERRMEEV-----YFDMAEDITLIFFHM-MGEPHETRHLVMHNEQAVISPMS---IHT	246
OY	327	GMEKNLTLMPRINGRIKESSSGSQPDNDPILDIYGYGDVR	368
Db	247	GVGTKNYAF---IWGM---GENLTFFD---MDHIAMDILDR	278
RESULT	12		
O00443			
ID	O00443	PRELIMINARY;	PRT; 1686 AA.
AC	O00443;		
DT	01-JUL-1997 (TREMBLrel_04, Created)		
DT	01-JAN-1998 (TREMBLrel_05, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel_17, Last annotation update)		
DE	PHOSPHOINOSTIDE 3-KINASE.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC	Mammalia; Eutheria; Primates; Catarrhini; Hominae; Homo.		
OX	NCBI_Taxid=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	MEDLINE=97479209; PubMed=9337861;		
RA	Domlin J., Pages F., Volinia S., Rittenhouse S.E., Zvelebil M.J.,		
RT	Stein R.C., Waterfield M.D.;		
RT	"Cloning of a human phosphoinositide 3-kinase with a C2 domain that		
RT	displays reduced sensitivity to the inhibitor wortmannin."		
RL	Biochem. J. 326:139-147(1997).		
DR	EMBL; Y13367; CAJ73797.1; -.		
DR	HSSP; P21707; 1BYN.		
DR	InterPro; IPR000008; C2.		
DR	InterPro; IPR001263; PI3Ka.		
DR	InterPro; IPR002420; PI3K_C2.		
DR	InterPro; IPR000341; PI3K_ras_bind.		
DR	InterPro; IPR000403; PI3_P14_kinase.		
DR	InterPro; IPR001683; PX.		
DR	pfam; PF00168; C2; 1.		
DR	pfam; PF00613; PI3Ka; 1.		
DR	pfam; PF00792; PI3K_C2; 1.		
DR	pfam; PF00794; PI3K_rbd; 1.		
DR	pfam; PF00434; PI3_P14_kinase; 1.		
DR	pfam; PF00787; PX; 1.		
DR	SMART; SM00239; C2; 2.		
DR	SMART; SM00145; PI3Ka; 1.		
DR	SMART; SM00146; PI3Kg; 1.		
DR	SMART; SM00142; PI3K_C2; 1.		
DR	SMART; SM00144; PI3K_rbd; 1.		
DR	SMART; SM00312; PX; 1.		
DR	PROSITE; PS50004; C2_DOMAIN_2; 1.		
DR	PROSITE; PS00915; PI3_A_KINASE_1; 1.		
DR	PROSITE; PS00916; PI3_A_KINASE_2; UNKNOWN_1.		
DR	PROSITE; PS50290; PI3_A_KINASE_3; 1.		
QC	SEQUENCE 1686 AA; E931IC803025C96F CRC64;		

Query Match	4.8%;	Score 114;	DB 4;	Length 1686;
Best Local Similarity	19.0%;	Pred. No. 0.92;		
Matches	85;	Conservative	68;	Mismatches 154;
				Indels 140;
				Gaps 21;

OY	45	ODNEPDIIV-----	OSAFQASASTYDANLDE-----	HEPELTYTALNCKTMILN-----	88
Db	522	EDDEPVDLNKHLQYIEKPCKEAMTRHPEVELLDSSYNQVFLAQINQHRANDQYKAV			581
OY	89	---CSALNODIMRLACYDVLVHGETPAVYIKTKRSIRIDETIMOTIKGPOVYVQETTDPI			145
Db	582	RKICSAID-GVETLAI-----	TESYKKRLRAVNLPIS-----	KTADVT	618
OY	146	FLMGNEKMLTKKAKOLEYAAKQFTPLSLSFDDDRNKTPLMSSRPINPMY-----	LP	199	
Db	619	SLFGE-----	-----DYSRSST-RGSLNPNPVQVSYINOLTAA	651	
OY	200	IF-----MHGKPNRSPNPTSHBARQFTPEFAPALPKQVSKVKAADL-----	WCTDSDLW	252	
Db	652	IYDLRLHLANSGRSFTDCQASSKYSKKEAMTTTEQLQFTIF-----	AAHGICSNMVSNEYK	707	
OY	253	FGYTOOSHWOJFNGKN-SRPF-----RVHDYQPEIFLTQ-----	PV-YSDELPMQKPRMI	300	
Db	708	YLICLSH-----	NCKDLFKPIQSKKVGYYKKNFFYLKMDLEIFPIQISQPLPESVHLT	766	
OY	301	GMGAVYHNSGESAKLSRSWN-----	RAYLAGMEWKNLTVPRWIGIIFK	345	
Db	764	LFGLINQSSGSSPDSNKKORKPEALGKYSLELDCDFRRFLTCG-----	TKLLYLW-----	812	
OY	346	EGSGGQPDNDPIDLYYGVGVPRFLYOLEKNSNISGVYRVPKSGKALQJLDYVYPLGK	405		
Db	813	---TSSHTNSVGTGYTKKGYVMERYLVQVDPSPAFDIIYTTTPVQDRSITIQOHNLTELEND	870		
OY	406	ISGYQFQFGYGSGLIDYNHEATSPGV	432		
Db	871	IKG-----	KLIDLHMDSSIGL	887	
RESULT	13				
ID	063485	PRELIMINARY:	PRN:	602 AA.	
AC	063485:				
DT	01-NOV-1996 (TREMBLrel. 01, Created)				
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)				
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)				
DE	C-RAP ACTIVATED ONCOGENE FUSION PROTEIN.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus				
OX	NCBI_TaxID=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=87172791; PubMed=3550433;				
RA	Ishikawa F., Takaku F., Nagao M., Sugimura T.;				
RT	"Rat c-rap oncogene activation by a rearrangement that produces a				
RT	fused protein."				
RL	Mol. Cell. Biol. 7:1226-1232(1987).				
CC	-1- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.				
DR	EMBL, M15428; AAA42002.1; -				
DR	InterPro: IPR000719; Euk_kinase.				
DR	InterPro: IPR002290; Ser_thr_kin_acctsite.				
DR	Pfam: PF00069; kinase.1.				
DR	SMART: SM00221; STKc.1.				
DR	PROSITE: PS00107; PROTEIN_KINASE_ATP.1.				
DR	PROSITE: PS50011; PROTEIN_KINASE_DOM.1.				
DR	PROSITE: PS01068; PROTEIN_KINASE_ST.1.				
KW	ATP-binding; kinase; Oncogene; Serine/threonine-protein kinase;				
KW	transferase.				
SO	SEQUENCE	602 AA;	69209 MW;	88B7BFA90FEB02AC CRC64;	

Query Match	4.78;	Score 110.5;	DB 11;	Length 602;
Best Local Similarity	19.58;	Pred. No. 0.43;		
Matches	93;	Conservative 68;	Mismatches 171;	Indels 145;
				Gaps 23

```

QY 29 NPAAVDEVRSENDLQODNLPIDVQASATSDTANPLDEHPELYTALENKTMLIN 88
D 87 NQVAKTEKKKELETAODRNLGI-----OSQFRAKEELEAEKRDLIRT---NERLSQ 137
QY 89 CSALNODIMRLACYDTLVHGETPAVITKRSIRLDETIMO--TFIGKPOVYVOE----- 140
D 138 VELTDEYKRL--NEKIKESNT---TKGELOKLDELQADYVYKYEKLEQEKELLN 192
QY 141 -----TTDPIFLMNGEK-----MLTKKAKOLEYAAKOFPLSLSPDL 180
D 193 ONSMLNTEKTKTDELLALGREGKNEILTELKCTLENKEDAIRSHSESASPALS--SSP 251
QY 181 RNTTPLMSSRPHNPMYLPFLFMHGKPNRSPTNSHEAROTPNFEFRAPELKFOVSAYVKA 240
D 252 NNLSPFGWSQPKTP-----VPAQREARAPSGTQEKKNKIRPGQORSYYWEIEASEVM 304
QY 241 AEDLMGTSDLMFGYTOOSH-----QIFNGKNSRPFVHDYOPEI----- 281
D 305 LSTRIGSGS---FGYVYKGNHGDVAVKILKYVDPTPEQLOAFNEVAVLKRTHVNL 361
QY 282 -----FLVQ-----PVYSDL--PMDQVYMI-----GMGAHVHNSGES 313
D 362 FMGYMTKDNLAIVYQWCEGSSLYKHLHVQETKFMQFOLIDIARQACQMDYLHAKNIIHR 421
QY 314 KLSRSMNRATLMAGMEK-----NLTVPRIMGRIFKESGSG----- 351
D 422 DMKS--NNIFLHGLVYKIGDFGLATVKSRL-----SSQVQOEPQSVLMAVAPYIR 472
QY 352 -PPDNPDIL--DYGYGDAVRFYOLENKSNIQSTVRYNPPSGKALQLDVYPLGKG 405
D 473 MQDNPPSPQSDVYSYIV--LYEL-----MTGELPYSHINNROI-----IFWYGRG 518

```

```

RESULT 14
ID 059241 PRELIMINARY: PRT: 821 AA.
AC 059241:
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE ENDO-1,4-BETA-GLUCANASE (EC 3.2.1.4) (CELLULOSE) (ENDOGALACTANASE)
OS CARBOXYMETHYL CELLULOSE.
OC Bacillus sp.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1409;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KSM-64;
RA Sumitomo N., Ozaki K., Ito S.;
RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ENDODIHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
LINKAGES IN CELLULOSE.
EMBL: M84963; AA073189.1; -
DR HSSP: 085465; 1A3H;
DR InterPro: IPR001347; Glyco_hydro_F5.
DR Pfam: PF00150; cellulase; 1.
DR PROSITE: PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN_1.
KW Hydrolase; Glycosidase.
SQ SEQUENCE 821 AA; 90910 MW; 73DA38FEF0B40B5C CRC64;

```

Query Match 4.6%; Score 109.5; DB 2; Length 821;
 Best Local Similarity 19.8%; Pred. No. 0.81;
 Matches 95; Conservative 54; Mismatches 163; Indels 169; Gaps 25;

```

QY 9 TLSTLSCFALIAIOAKAVNPVAFVDEVRSEND---LQODNELPIDVQASATSDTA 65
D 6 TKOLISSILIVL-LSLFPALAAEGTREDNKKHLGNDNVKRPSEAGALQLOEDVQ 64
QY 66 NPL-DEHEPELY-----TTALENKTMLINCSAL-----NODIMRLACYDTLVHGETP 111

```

```

D 65 MFLVDQGEKIQLRGMSHGLQWPEPELNDNAAYKALANDMSNMIRLAMYGENGYASN 124
QY 112 AVITKRSIRLDETI-----WQT-----IKKPOVY 138
D 125 ELIKSRVYKIGIDLAIENDMYIVDMVHAPDPRDPVYAGAEDEFRLIAYLPNNPHIY 184
QY 139 QETDP-----IFLNGNEKMLTKKDAKOLEYAAKOFPLSLSPDLR-----NPLW 187
D 185 ELANEPSSNNNGAGCIPNNEGMNAV-----ETADPIVEMLRSGANDNDIIVGSPNW 239
QY 188 SSRP-----HNPVYLPPIF--MHGKPNRS--PNTPSHEAROTPNFEFRAPELK 233
D 240 SQRPDLADNPIDDHHTMYTHFTGSHAASTESPPEPTPSEGNVNSNRYA-----LE 295
QY 234 VSVKKAEDLMGT-----DSDLMFGYTOO-----SHWQIFNGKNS-----RP 271
D 296 NGVAVFATE--WGTQANGDGCPYFEDADVIEEFLNENNISMAWMSLTN--KNEVSGATP 352
QY 272 FRVH-----DYQPEIFLQPVYSDDLPM---D 294
D 353 FELKSNATSLDPPDDOYVPEEUSLGEYVRARIKGVNPE---IDRTKTYKVLMDPND 409
QY 295 GKVMIGMGAHVHNSGES-----AKLSRSMNRATLMAGMEKMLTVAPRTWG 341
D 410 GTKGFGV-----NGDSPVEDVYIENAGALKISGLDASNDVSGCNWAMARLSADG 463
QY 342 R 342
D 464 K 464

```

```

RESULT 15
ID 09F216 PRELIMINARY: PRT: 824 AA.
AC 09F216:
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE CELLULOSE.
OS Bacillus sp.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1409;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KSM-6237;
RC MEDLINE=21036886; PubMed=11193393;
RA Hakamada Y., Hatada Y., Koike K., Yoshimatsu T., Kawai S.,
RA Kobayashi T., Ito S.;
RT "deduced amino acid sequence and possible catalytic residues of a
thermostable, alkaline cellulase from an alkaliphilic Bacillus
strain.";
RL Biosci. Biotechnol. Biochem. 64:2281-2289(2000).
DR EMBL: AB018420; BAB19360.1; -
DR InterPro: IPR001547; Glyco_hydro_F5.
DR Pfam: PF00150; cellulase; 1.
DR PROSITE: PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN_1.
SQ SEQUENCE 824 AA; 91564 MW; 65FA940FE1D729B9 CRC64;

```

Query Match 4.6%; Score 108; DB 2; Length 824;
 Best Local Similarity 19.8%; Pred. No. 1.1;
 Matches 95; Conservative 60; Mismatches 159; Indels 166; Gaps 28;

```

QY 9 TLSTLSCFALIAIOAKAVNPVAFVDEVRSEND---LQODNELPIDVQASATSDTA 65
D 7 TKOLISSILIVL-LSLFPALAAEGTREDNKKHLGNDNVKRPSEAGALQLOEDVQ 65
QY 66 NPL-DEHEPELY-----TTALENKTMLINCSAL-----NODIMRLACYDTLVHGE-- 109
D 66 MFLVDQGEKIQLRGMSHGLQWPEPELNDNAAYKALANDMSNMIRLAMY-----VGENGY 121
QY 110 --TVAIVTK-----RSIRLDETI--WQT-----IKKRP 134

```

```
Db 122 ANPELIKORVIDGIELALENDMYIVDMHVAHPGDPDPVYAGAKDFPREIALALYPNNP 181
QY 135 QVVOETTP-----IFLGNKEKMLTKDAKOLEYAKOFTPLSLFDDLRN-----N 183
Db 182 HITTELANEPSNNNGAGIPNNEGMAK-----ETADPIYEMLRKSGNADDDNIIIVG 236
QY 184 TPLWSSRP-----HNPVYLPFIMHGKPNRSPNTPSHEARQFTPNFRAPELK-- 231.
Db 237 SPNMSQRPDLADNPIDHHTMYTVHFYTGSHAASJESYPS-----TPNSERGNVMSNT 291
QY 232 ---FOYSVKVKAEDLMGT-----DSDLMFCTTQQ-----SHWQIFNGKNS-- 269
Db 292 RYALENGAVAFATE--WGTSQASGDGPYFDEADWIEFLNENNISMAMNSITN-KNEYS 348
QY 270 ---RPFV-----HDYOP-EIFLT-----QPV-----YSDLPW-- 293
Db 349 GAFTEPELCKSNATNIDPGPDHVAPEELSLSGEYVRARIKGVNYEPIDRTKTKYLMDF 408
QY 294 -DGKVRMIGMGAVHHSNGESAKLSRSWNPAYLWAGME-----WKNLTVMPRIWGR 342
Db 409 NDGTQO--GEGVNSDSPNKELIAVDNENNTLKVSGLDVSNDSGNFMANARLSANGWGK 466
```

Search completed: November 30, 2001, 14:26:28
Job time: 565 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:27:03 ; Search time 28.02 Seconds

(without alignments)
578.367 Million cell updates/sec

Title: US-09-787-083-6
Perfect score: 2360
Sequence: 1 MKVSLSTLTSLSCFAILA.....YNHATSFGVGLMNDWMGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues
Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	373	15.8	286	1	PAL_KLEPN
2	358	15.2	289	1	P37446 klebsiella
3	357	15.1	289	1	PAL_SALTY
4	321.5	13.6	289	1	PAL_ECOLI
5	109.5	4.6	800	1	PAL_PROVU
6	98.5	4.2	810	1	GUN_BACS1
7	97.5	4.1	1478	1	AMDH_YEAST
8	95	4.0	719	1	BCK1_YEAST
9	94.5	4.0	397	1	PBPA_STRPN
10	94	4.0	716	1	GUN_PAPEO
11	94	4.0	1115	1	BAC2_MOUSE
12	92.5	3.9	849	1	DR3A_BACSU
13	92.5	3.9	849	1	SRK6_BRAOL
14	92.5	3.9	1379	1	SRK6_BRAOL
15	92	3.9	324	1	MET_MOUSE
16	92	3.9	888	1	YJHO_ARCFU
17	91.5	3.9	486	1	YJHO_YEAST
18	91.5	3.9	969	1	YJHO_YEAST
19	91	3.9	969	1	SACB_STRSL
20	90.5	3.8	496	1	CATA_DICDI
21	90.5	3.8	1024	1	INVO_MOUSE
22	90	3.8	1158	1	Y075_MYCGE
23	90	3.8	1158	1	Y075_MYCGE
24	89	3.8	1788	1	R114_HUMAN
25	88.5	3.8	825	1	NP72_CAEEL
26	88.5	3.8	825	1	NP72_CAEEL
27	88	3.7	1382	1	NP72_CAEEL
28	88	3.7	1382	1	NP72_CAEEL
29	87.5	3.7	478	1	DRHG_AICCA
30	87.5	3.7	503	1	CD44_RAT
31	87.5	3.7	669	1	CD44_RAT
32	87	3.7	637	1	AMY_ALPHA
33	87	3.7	666	1	PBPA_STROR

34	87	3.7	828	1	BGAL_BRAOL	P49676 brassica ol
35	87	3.7	842	1	LPFC_SALTY	P43662 salmonella
36	87	3.7	917	1	YJ13_YEAST	P53148 saccharomyc
37	87	3.7	1474	1	A2MG_HUMAN	P01023 homo sapien
38	86	3.6	1008	1	SN14_YEAST	P36048 saccharomyc
39	86	3.6	1087	1	XYNX_CLOTH	P38535 clostridium
40	85.5	3.6	353	1	DCUP_BACSU	P32395 bacillus su
41	85.5	3.6	790	1	SEIL_MOUSE	O92266 mus musculu
42	85.5	3.6	794	1	SEIL_MOUSE	O92266 mus musculu
43	85.5	3.6	1374	1	YC9A_SCHPO	O09884 schizosacch
44	85	3.6	491	1	TY3H_PNASP	P11962 phasianidae
45	85	3.6	525	1	MP11_RAT	P48965 rattus norv

ALIGNMENTS

RESULT 1	ID	Pal_KLEPN	STANDARD	PRT	286 AA.
AC	P37446				
DT	01-OCT-1994 (Rel. 30, Created)				
DT	01-OCT-1994 (Rel. 30, Last sequence update)				
DT	01-NOV-1997 (Rel. 35, Last annotation update)				
DE	PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT				
DE	PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-				
DE	ACTYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).				
GN	PLDA.				
OS	Klebsiella pneumoniae.				
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;				
OC	Klebsiella.				
OX	NCBI_TaxID=573;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=94131966; PubMed=8300539;				
RA	Brok R.G.P.M., Brinkman E., van Boxtel R., Bekkers A.C.A.P.,				
RA	Verheij H.M., Tomassen J.;				
RT	"Molecular characterization of enterobacterial plda genes encoding				
RT	outer membrane phospholipase A.";				
RL	J. Bacteriol. 176:861-870(1994).				
CC	- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE				
CC	A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.				
CC	- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-				
CC	PHOSPHOCHOLINE + A FATTY ACID ANION.				
CC	- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-				
CC	PHOSPHOCHOLINE + A FATTY ACID ANION.				
CC	- Cofactor: REQUIRES CALCIUM IONS FOR ACTIVITY.				
CC	- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES				
CC	LOCATED THERE.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; X76901; CAA54223.1; -				
DR	PIR; B36971; B36971.				
DR	PIR; S40129; S40129.				
DR	InterPro: IPR003187; PLA1.				
DR	Pfam: PF02253; PLA1; 1.				
KW	Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium.				
FT	STGNAL				
FT	CHAIN				
FT	ACT SITE				
FT	SEQUENCE				

Query Match 15.8%; Score 373; DB 1; Length 286;
Best Local Similarity 42.1%; Pred. No. 5,6e-23;
Matches 90; Conservative 27; Mismatches 87; Indels 10; Gaps 5;

```
QY 226 RAPELKFQVSVKVAEDLWGTDSDLMFGYTQOOSHMOJFNGKNSRPFVHYDQPELFLQ 285
| :|||: : : :|||: : : :|||: : : :|||: : : :|||: : : :|||:
DB 80 RAEVYKFOJSLAFPLMRGILGDNLSLGASYTQKSMWQJNSKESAPFRTNTEPQLGLGF 139
QY 266 PV-YSDLPWDGKVRNIGMGAVHNSGESAKLSRSNNRAYLAGMEKKULTVPRIMGRTF 344
| :|||: : : :|||: : : :|||: : : :|||: : : :|||: : : :|||:
DB 140 ATDYQFAGW--TLRDIEKGYNHDNSGRSDPTSRNNRLYLRLMAONGMMLVEKPM--Y 194
QY 345 KEGSGSQPDNDPDLIDYGYGDFRFLYOLENKSNSGTVRYNPRSGKALQLDYVYPLGK 404
| :|||: : : :|||: : : :|||: : : :|||: : : :|||: : : :|||:
DB 135 VVG---TDDNPDITRYKMYRLKGYQL-GEALISAGQGYMNMWNGYGAELGVSTPIRK 250
QY 405 GISGYFOIFQGYGOSLIDYNHEATSEFGVGLMND 438
| :|||: : : :|||: : : :|||: : : :|||: : : :|||: : : :|||:
DB 251 HVRYATQIYSGYGESLIDYNFNQTRVGVGLMND 284

RESULT 2
PAL_SALTY STANDARD: PRT: 289 AA.
ID PAL_SALTY
AC P37442:
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT
DE PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-
DE ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).
GN PLDA.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC NCBI_TaxID=602;
OX [1]
RN
RP MEDLINE=94131966; PubMed=8300539;
RA Brok R.G.P.M., Brinkman E., van Bortel R., Bekkers A.C.A.P.,
RT Verheij H.M., Tommassen J.;
RT "Molecular characterization of enterobacterial plid genes encoding
RT outer membrane phospholipase A."
RL J. Bacteriol. 176:861-870(1994).
CC -1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE
CC A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-
CC -1- PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-
CC -1- PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.
CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES
CC LOCATED THERE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: X76900; CA54222.1; -
DR PIR: A36971; A36971.
DR PIR: S40131; S40131.
DR StyGene: SG10302; PLD1.
DR InterPro: IPR003187; PLD1.
DR Pfam: PF02253; PLD1; 1.
KW Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium.
FT SIGNAL 1 20 BY SIMILARITY.
FT CHAIN 21 289 PHOSPHOLIPASE A1.
FT ACT_SITE 164 164 BY SIMILARITY.
SQ SEQUENCE 289 AA; 33003 MW; D5ACF5E33F704DBA CRC64;
```

Query Match

15.2%; Score 358; DB 1; Length 289;

```
Best Local Similarity 39.3%; Pred. No. 9.2e-22;
Matches 84; Conservative 30; Mismatches 90; Indels 10; Gaps 4;
QY 226 RAPELKFQVSVKVAEDLWGTDSDLMFGYTQOOSHMOJFNGKNSRPFVHYDQPELFLQ 285
| :|||: : : :|||: : : :|||: : : :|||: : : :|||: : : :|||:
DB 80 RAEVYKFOJSLAFPLMRGILGDNLSLGASYTQKSMWQJNSKESAPFRTNTEPQLGLGF 142
QY 266 PV-YSDLPWDGKVRNIGMGAVHNSGESAKLSRSNNRAYLAGMEKKULTVPRIMGRTF 344
| :|||: : : :|||: : : :|||: : : :|||: : : :|||: : : :|||:
DB 140 ATDYQFAGW--TLRDIEKGYNHDNSGRSDPTSRNNRLYLRLMAONGMMLVEKPM--Y 199
QY 345 KEGSGSQPDNDPDLIDYGYGDFRFLYOLENKSNSGTVRYNPRSGKALQLDYVYPLGK 404
| :|||: : : :|||: : : :|||: : : :|||: : : :|||: : : :|||:
DB 200 -----GSTDDNDPDIKRYMGYQLKIGYHL-GEAVLSAKQGYMNMWNGYGAELGVSTPIRK 253
QY 405 GISGYFOIFQGYGOSLIDYNHEATSEFGVGLMND 438
| :|||: : : :|||: : : :|||: : : :|||: : : :|||: : : :|||:
DB 251 HVRYATQIYSGYGESLIDYNFNQTRVGVGLMND 287

RESULT 3
PAL_ECOLI STANDARD: PRT: 289 AA.
ID PAL_ECOLI
AC P00631:
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT
DE PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-
DE ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).
GN PLDA OR B3821 OR Z5342 OR ECS4751.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC NCBI_TaxID=562, 83334;
OX [1]
RN
RP MEDLINE=85157492; PubMed=6397464;
RA Homma H., Kobayashi T., Chiba N., Karasawa K., Mizushima H., Kudo I.,
RT Inoue K., Ikeda H., Sekiguchi M., Nojima S.;
RT "The DNA sequence encoding plid gene, the structural gene for
RT detergent-resistant phospholipase A of E. coli."
RL J. Biochem. 96:1655-1664(1984).
CC [2]
CC SEQUENCE FROM N.A.
CC STRAIN=K12 / MG1655;
CC MEDLINE=92358234; PubMed=1379743;
CC Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;
CC "Analysis of the Escherichia coli genome: DNA sequence of the region
CC from 84.5 to 86.5 minutes."
CC Science 257:771-778(1992).
CC [3]
CC REVISION TO 14-15.
CC STRAIN=K12 / MG1655;
CC MEDLINE=97426617; PubMed=9278503;
CC Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
CC Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
CC Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
CC Mau B., Shao Y.;
CC "The complete genome sequence of Escherichia coli K-12."
CC Science 277:1453-1474(1997).
CC [4]
CC SEQUENCE FROM N.A.
CC STRAIN=O157:H7 / EDL933 / ATCC 700927;
CC MEDLINE=21074935; PubMed=11206551;
CC Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
CC Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
CC Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
CC Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
CC Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
CC Welch R.A., Blattner F.R.;
CC "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
CC
```


OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90105403; PubMed=2690949;
 RA Meyer S.L., Kvalnes-Krick K.L., Schramm V.L.:
 RT "Characterization of AMD, the AMP deaminase gene in yeast. Production
 of amd strain, cloning, nucleotide sequence, and properties of the
 RT protein."
 RL Biochemistry 28:8734-8743(1989).
 (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=5288C / AB972;
 RA Badcock K., Churcher C., Barrell B.G., Rajandream M.A., Walsh S.V.:
 RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY
 CC METABOLISM.
 CC -1- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).
 CC -1- PATHWAY: PURINE NUCLEOTIDE CYCLE.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SIMILARITY: BELONGS TO THE ADENOSINE AND AMP DEAMINASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M30449; AAA34420.1; -
 DR EMBL: Z46659; CAA86620.1; -
 DR PIR: A33365; A33365.
 DR SGD: S0004498; AMD1.
 DR InterPro: IPR001365; A_deaminase.
 DR Pfam: PF00962; A_deaminase.1.
 DR PROSITE: PS00485; A_DEAMINASE.1.
 DR HydroLase: Nucleotide metabolism.
 FT ACT_SITE 422 422 POTENTIAL.
 FT ACT_SITE 631 631 POTENTIAL.
 FT ACT_SITE 707 707 POTENTIAL.
 FT ACT_SITE 708 708 POTENTIAL.
 FT CONFLICT 568 568 F -> C (IN REF. 1).
 SQ SEQUENCE 810 AA; 93301 MW; 7A6DCB439B45C93 CRC64;

Query Match 4.2%; Score 98.5; DB 1; Length 810;
 Best Local Similarity 21.3%; Pred. No. 3.7;
 Matches 74; Conservative 39; Mismatches 132; Indels 103; Gaps 15;

QY 46 DNELPDIYQASATQASTDPA-----PLDEHEPELYTTALENTML 86
 DB 48 DETPPEODDSHESLADSNANFSYHENOQLLENGTKOLALDEHOS--HSALTEPSHS 105
 QY 87 INCNALNODIMRLACYDTLVH-----GETPAVYIKTRSRILDETIMOTIKGPQVYVYQET 141
 DB 106 TNCSSSNIAAMNKG-HDSADHASQNSGKPRITLSASAQHLPELT- KSRFGAIVVVKQVR 163
 QY 142 TDBPIFLMG-----NEKGMILTKDAKOLEYAAKQFTPLISLFDLRNNTPL 186
 DB 164 TSASYKMGMLADDAISOQLDDPSSSELIDLYSKVAECRNLRKAKYQITISVQNDQNPKNKPG 223
 QY 187 W--SSRPNHPMY-----VLPIFMHGKPNRSPPTPSHEARQFPNPFRAELKFQVSVKV 238
 DB 224 WVVYPPPKPSTNSDKTKTVV-----TKKPAEYFDFKCE----- 260
 QY 239 KAEDLACGTDSDLMFGYTOQSHWQIFNGKNSRPFVHDYQPEFLTPQVYSDLPMDGKVR 298
 DB 261 -----IPGEDPDEWFLTNDSDYV-----HRSGKTDDELTAQIPLTLDYLDLDE-----K 305
 QY 299 MIGMGAVHNSGESAR-----LSRSWNRAYLAKGMEKNTLVMPR 338
 DB 306 MISIS-----SDGPAKSFAYRRLQYLEARNMLLYLL--NEYQETSVSKR 347

RESULT 7
 ID BCK1_YEAST STANDARD; PRT: 1478 AA.
 AC 001389; P32894;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE SERINE/THREONINE PROTEIN KINASE BCK1/SLK1/SSP31 (EC 2.7.-.-).
 GN BCK1 OR SLK1 OR SSP31 OR LAS3 OR YJL095W OR J0906.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92186847; PubMed=1545797;
 RA Costigan C., Gehring S., Snyder M.;
 RT "A synthetic lethal screen identifies SLK1, a novel protein kinase
 RT homolog implicated in yeast cell morphogenesis and cell growth."
 RL Mol. Cell. Biol. 12:1162-1178(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92104496; PubMed=1840547;
 RA Irie K., Araki H., Oshima Y.;
 RT "A new protein kinase, SSP31, modulating the SMP3 gene-product
 RT involved in plasmid maintenance in Saccharomyces cerevisiae."
 RL Gene 108:139-144(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EG123;
 RX MEDLINE=92107166; PubMed=1729597;
 RA Lee K.S., Levin D.E.;
 RT "Dominant mutations in a gene encoding a putative protein kinase
 RT (BCK1) bypass the requirement for a Saccharomyces cerevisiae protein
 RT kinase C homolog."
 RL Mol. Cell. Biol. 12:172-182(1992).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=5288C;
 RX MEDLINE=95116706; PubMed=7871887;
 RA Miosga T., Boles E., Schaeff-Gerstenschlaeger I., Schmitt S.,
 RA Zimmermann F.K.;
 RT "Sequence and function analysis of a 9.74 kb fragment of
 RT Saccharomyces cerevisiae chromosome X including the BCK1 gene."
 RL Yeast 10:1481-1488(1994).
 RN [5]
 RP SEQUENCE OF 602-1104 FROM N.A.
 RA Cusick M.E.;
 RL Submitted (XX-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: SERINE/THREONINE PROTEIN KINASE INVOLVED IN A SIGNAL
 CC TRANSDUCTION PATHWAY THAT PLAY A ROLE IN YEAST CELL MORPHOGENESIS
 CC AND CELL GROWTH. THIS PATHWAY SEEMS TO STARTS BY SMP3; THEN
 CC INVOLVE THE KINASE PC1 THAT MAY ACT ON THIS KINASE. BCK1 PROBABLY
 CC PHOSPHORYLATES MK1 AND MK2 WHICH THEMSELVES PHOSPHORYLATE THE
 CC MPX1 KINASE.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE KINASE KINASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M84389; -; NOT_ANNOTATED_CDS.
 DR EMBL: D10389; BAA01226.1; -
 DR EMBL: X60227; CAA42788.1; -
 DR EMBL: X77923; CAA54896.1; -
 DR EMBL: Z49370; CAA89389.1; -

DR EMBL; Z49369; CA89388.1; -;
 DR EMBL; M88604; AAA21179.1; -;
 DR PIR; S20117; S20117.
 DR PIR; S22285; S22285.
 DR PIR; J01118; J01118.
 DR PIR; J01432; J01432.
 DR HSP; P24941; IAO1.
 DR SGD; S0003631; BCK1.
 DR InterPro; IPR000719; Euk_Pkinase.
 DR InterPro; IPR002290; Ser_thr_kin_actsite.
 DR Pfam; PF00069; pkinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR Kinase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
 KW ATP-binding; Phosphorylation.
 FT DOMAIN 1175 1440 PROTEIN KINASE.
 FT NP_BIND 1181 1189 ATP (BY SIMILARITY).
 FT BINDING 1204 1204 ATP (BY SIMILARITY).
 FT ACT_SITE 1303 1303 BY SIMILARITY.
 FT MOD_RES 1134 1134 PHOSPHORYLATION (BY PKC) (POTENTIAL).
 FT VARIANT 1119 1119 T -> P (IN BCK1-19; ACTIVATION).
 FT VARIANT 1120 1120 I -> K (IN BCK1-11; ACTIVATION).
 FT VARIANT 1120 1120 I -> T (IN BCK1-16; ACTIVATION).
 FT VARIANT 1146 1146 G -> V (IN BCK1-10; ACTIVATION).
 FT VARIANT 1174 1174 A -> P (IN BCK1-20; ACTIVATION).
 FT CONFLICT 59 59 E -> I (IN REF. 2).
 FT CONFLICT 79 79 E -> V (IN REF. 3).
 FT CONFLICT 264 264 A -> P (IN REF. 3).
 FT CONFLICT 279 279 N -> I (IN REF. 3).
 FT CONFLICT 703 714 RYPTSTYIDR -> STPKRVTMT (IN REF. 3).
 FT CONFLICT 795 795 S -> A (IN REF. 3).
 FT CONFLICT 802 802 L -> V (IN REF. 3).
 FT CONFLICT 808 808 A -> S (IN REF. 3).
 FT CONFLICT 903 903 T -> N (IN REF. 3).
 FT CONFLICT 919 919 T -> N (IN REF. 3).
 FT CONFLICT 960 962 ADA -> RDR (IN REF. 5).
 FT CONFLICT 1086 1104 RPPVDSSEYFIQDLNGK -> VPIAHTSSYRMDLVKIN
 H (IN REF. 5).
 SQ SEQUENCE 1478 AA; 164194 MW; D586C3A497A5BB3 CRC64;

Query Match 4.1%; Score 97.5; DB 1; Length 1478;
 Best Local Similarity 23.0%; Pred. No. 10;
 Matches 110; Conservative 50; Mismatches 175; Indels 143; Gaps 26;

QY 2 KVSITSLTSL-----ILSCFAILAIQOAKAVPNPFAFVDEVRESENDLQDNELPT---DV 53
 DB 912 KVNNSNSTVSTNSITFTSPSPFLKRGSKRVVSTSAAD-ITFEENDITPADPPEFSDSD 970
 QY 54 QSAIQSASTDTPANPLDEHEPELYTTALENK--TMDLINCASALNODIMRLACYDTLVHGEP 111
 DB 971 SDDSSSSDDIDMSKKKAPE---TNNENKKDEKSDNSTHSELF-----YDS---QTO 1019
 QY 112 AVITKRSIRLDEITWQITKGPQVYQETDPTFLMGNEKGMILTAKAKOLEVAAKQFT 171
 DB 1020 DKMERK-----MTFRSPPEVYVO-NLEKFPFRANLDPKPT-----EGIASPTS 1061
 QY 172 PLTSLFDDLRNNT-----PLWSSRPHPMYVLPFIMHG-----KPNSSPMP----- 213
 DB 1062 PKSLDLSLSPKNAVSSRIEPTSPRPVPPDSSYEFTIQDLGKKNKPLMOAKTPKTKTIR 1121
 QY 214 --SHEARFTNEFRAPELKQVSVKAAE-DLWGTDSDLWFGYTOOSHMOIFNGKNSR 270
 DB 1122 TIAHSAIARKN-----SVKLKRONTKMGT---RMVEVTEHHMVSINNAKNSK 1167
 QY 271 PPRVHDYQPELFTLPQVYSDLEWDSKVMIGCA-----VHNSGESAKISR----- 317
 DB 1168 -----GEYKEFAW-MKGMETIKGSGFAYVLCINVTGEMAMVAKQVEVPKY 1211
 QY 318 -SWNRAVLMAGMEKMLVMPRIWGRIFKEGSGSPDPNDPLDLYYGDAVFYOLENK 376
 DB 318 -SWNRAVLMAGMEKMLVMPRIWGRIFKEGSGSPDPNDPLDLYYGDAVFYOLENK 376

DB 1212 SSQNEALIS-----TV-----EALRSEVSTLKDLHLNIYVYLG-----ENK 1249
 QY 377 SNISGTVRNPRSGCALQLDVYVPLGKISGYFOIFGCGSLDYNHREATSPGVGL 434
 DB 1250 NNTY-----SIFLEYV--AGGSVGSILRMGRFDEPLI--KHLLTYVLKGL 1291
 RESULT 8
 ID BPBA_STRPN STANDARD: PRT: 719 AA.
 AC 004707;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A (PBP-1A) (EXPORTED PROTEIN 2).
 GN PONA OR EXP2.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacilllus/Clostridium group; Streptococcaceae;
 CC Streptococcus.
 OX NCBI_TaxID=1313;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=45607, AND 63915;
 RC MEDLINE=93010977; PubMed=1396576;
 RA Martin C., Sibold C., Hakenbeck R.;
 RT "Relatedness of penicillin-binding protein 1a genes from different
 clones of penicillin-resistant Streptococcus pneumoniae isolated in
 South Africa and Spain."
 RL EMBO J. 11:3831-3836(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92325042; PubMed=1624444;
 RA Martin C., Briese T., Hakenbeck R.;
 RT "Nucleotide sequences of genes encoding penicillin-binding proteins
 from Streptococcus pneumoniae and Streptococcus oralis with high
 homology to Escherichia coli penicillin-binding proteins 1a and 1b."
 RL J. Bacteriol. 174:4517-4523(1992).
 RN [3]
 RP SEQUENCE OF 293-369 FROM N.A.
 RC STRAIN=RX6;
 RX MEDLINE=95020625; PubMed=7934910;
 RA Pearce B.J., Yin Y.B., Masure H.R.;
 RT "Genetic identification of exported proteins in Streptococcus
 pneumoniae."
 RL Mol. Microbiol. 9:1037-1050(1993).
 CC -1- FUNCTION: CELL WALL FORMATION.
 CC -1- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- MISCELLANEOUS: THESE STRAINS ARE PENICILLIN-SENSITIVE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X67873; CA448073.1; -;
 DR EMBL; X67872; CA448072.1; -;
 DR EMBL; M80527; AAA26956.1; -;
 DR PIR; S28038; S28038.
 DR InterPro; IPR001264; Transglycosyl.
 DR InterPro; IPR001460; Transpeptidase.
 DR Pfam; PF00912; Transglycosyl; 1.
 DR Pfam; PF00905; Transpeptidase; 1.
 DR ProDom; PD001895; Transglycosyl; 1.
 KW Peptidoglycan synthesis; Antibiotic resistance; Cell wall;
 KW Multifunctional enzyme.
 FT ACT_SITE 370 370 ACYLATED BY PENICILLIN (BY SIMILARITY).
 FT DOMAIN 657 683 SER-RICH.
 FT VARIANT 124 124 T -> A (IN STRAIN R6).
 FT VARIANT 386 386 V -> I (IN STRAIN 63915).

FT VARIANT 388 388 D -> E (IN STRAIN R6).
 FT VARIANT 397 397 E -> K (IN STRAIN 63915).
 FT VARIANT 523 523 M -> I (IN STRAIN 63915).
 FT VARIANT 533 533 E -> D (IN STRAINS 63915 AND R6).
 FT VARIANT 540 540 T -> S (IN STRAINS 63915 AND R6).
 FT VARIANT 657 657 S -> N (IN STRAIN 63915).
 SO SEQUENCE 719 AA; 79745 MW; 5BC3C6A93BFA970A CRC64;

Query Match 4.0%; Score 95; DB 1; Length 719;
 Best Local Similarity 20.1%; Pred. No. 6;
 Matches 96; Conservative 70; Mismatches 196; Indels 116; Gaps 25;

QY 3 VSLSTLTLSI-----LSCAIIAIIQAQKAVNPVAVDVRENDL 43
 DB 13 LLSLSFLVIAIIVLGGVFFYYVSKAPSLSESKLITVSSKIYDKNQLIADLGSRRV 72
 QY 44 -GODNELPIDVOSATOS-----ASTDTANPLDEHPELYTTALENKTMILNCSALN 93
 DB 73 NQAANDIPDLVKAIYISIEDHREFDRGIDITILIGAFILNLSNLSLGG-----STLT 126
 QY 94 QDIMRLACYDTLVHGETPVAIKTKRSIRLDETIMQTIKGPVYVQET---TDPIFLMG 149
 DB 127 QOLIKLTFESTSTSDQITIS-----RKAQEAMLAIQLQKATKQEIILTYINKYMSN 178
 QY 150 NKGMLTK-----KRAKOLE-----YAKOFTPLS-LSTDLDNRNTPLMSSRP 191
 DB 179 GNYGMOTAAQNYGKDLNLSLPOLALLAGMPQAPQDYDPSHPEAODRRNLVL--SEM 236
 QY 192 HNPVWYLPFIEMHGKPNRSPPTPSHEAROFPTNFRAPPE-----LKRQVS-VKVAEDLM 245
 DB 237 KNGGYI-----SAQYKRAVNTPTLDGLQSLKSNTPAYMDNLKEVINVEETGYNL 292
 QY 246 GTDSDLMFSTYTO--QSH-MQIFNGKNSRPRVHDYQPELFTLPVYSDLPMGDKVMIGM 302
 DB 293 TGTGMVYTVNDDEAKHLMWDIYNTDEVAVYPPDELO-----VASTIVDVS-NGKV-IAQL 345
 QY 303 GAVHSHN-----GESAKLSRSNNRAYLLAGMKNLT-VMPII-----WRIKREGS 348
 DB 346 GAHQSSNVSFGINQAVETNRDW-----GSTMKPTITVAPALEXYDSTATIYHDEP 398
 QY 349 GSPQDNPDLIDY-YGY-GDVRFYLOENKSNISGTVRYN-----PRSGKALOLDY 398
 DB 399 YNPGINTVYVYNDKRGYFGNITLQYALQOSRNVPAVETLKNKVLNRAKFTLNLGLGIDY 456

RESULT 9
 GUN_PAERO STANDARD; PRT: 397 AA.

AC P23548;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 RT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE ENDOGLUCANASE (EC 3.2.1.4) (ENDO-1,4-BETA-GLUCANASE) (CELLULASE).
 OS Paenibacillus polymyxa (Bacillus polymyxa).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Paenibacillus.
 OX NCBI_TaxID=1406;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90170877; PubMed=2307659;
 RA Baird S.D., Johnson D.A., Seligy V.L.;
 RT "Molecular cloning, expression, and characterization of
 RT endo-beta-1,4-glucanase genes from Bacillus polymyxa and Bacillus
 RT circularis";
 RL J. Bacteriol. 172:1576-1586(1990).
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
 CC LINKAGES IN CELLULOSE.
 CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY A (FAMILY 5 OF GLYCOSYL
 CC HYDROLASES).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M33791; AAA22631.1; -.
 DR PIR: A35136; A35136.
 DR HSSP: P54383; 1ECE.
 DR InterPro: IPR001547; Glyco_hydro_F5.
 DR Pfam: PF00150; cellulase_1.
 DR PROSITE: PS00659; GLYCOSYL_HYDROL_F5; 1.
 KW Cellulose degradation; Hydrolase; Glycosidase.
 FT ACT_SITE 194 194 PROTON DONOR (BY SIMILARITY).
 FT ACT_SITE 317 317 NUCLEOPHILE (BY SIMILARITY).
 SO SEQUENCE 397 AA; 44357 MW; B9C2E802C04F0A2A CRC64;

Query Match 4.0%; Score 94.5; DB 1; Length 397;
 Best Local Similarity 25.7%; Pred. No. 2.9;
 Matches 46; Conservative 21; Mismatches 61; Indels 51; Gaps 10;

QY 301 GMGAVHSHNGESAKLSRNNRAYLLAGMKNLTVMPIRIGR-----IFKEG----- 347
 DB 43 GNKIVDESGKEAFNGLNW-----FGLRPPNT-LHGLMSRSDMDMLDYKKEGYLLIR 95
 QY 348 -----SGSQPDD-----NPDLIDYGYGDVRFYLOENKSNISGTV-----RYNPRS 369
 DB 96 LPYSNQLFDSSSRPDSIDYHKNPDLV---GLNPQLQIMDKLIERAKGORGIOIILDRHPGS 152
 QY 390 GKALOLDYV--YPLKGIGISGYFOIFGY-----GQSLIDNNHETSGVGLMLNDW 439
 DB 153 G-GQSELWYTSQYPSRWISDMKMLADRYKNNPVIGADLHNEPHGQASWGTSNASTDW 210

RESULT 10

BAC2_MOUSE STANDARD; PRT: 716 AA.

AC P97303;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE TRANSCRIPTION REGULATOR PROTEIN BACH2 (BFB AND CNC HOMOLOG 2).
 GN BACH2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C;
 RX MEDLINE=97042438; PubMed=8887638;
 RA Oyake T., Itoh K., Motohashi H., Hayashi N., Hoshino H., Nishizawa M.,
 RA Yamamoto M., Igarashi K.;
 RT "Bach proteins belong to a novel family of BFB-basic leucine zipper
 RT transcription factors that interact with Mafk and regulate
 RT transcription through the NF-E2 site";
 RL Mol. Cell. Biol. 16:6083-6095(1996).
 CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS AS REPRESSOR OR
 CC ACTIVATOR. BINDS, IN-VITRO, TO NF-E2 BINDING SITES. PLAY IMPORTANT
 CC ROLES IN COORDINATING TRANSCRIPTION ACTIVATION AND REPRESSION BY
 CC MAFK.
 CC -1- SUBUNIT: HETERODIMER OF BACH2 AND MAFK.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSION RESTRICTED TO MONOCYTES AND
 CC NEURONAL CELLS.
 CC -1- SIMILARITY: BELONGS TO THE BZIP FAMILY. CNC SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 BFB/POZ DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: D86604; BAI13138.1; -
DR HSSP: P05412; IPOS.
DR MCD: MGI:894679; Bach2.
DR InterPro: IPR000210; BTB_POZ.
DR InterPro: IPR001871; bZIP.
DR Pfam: PF00651; BTB; 1.
DR Pfam: PF00170; bZIP; 1.
DR SMART: SM00338; BRLZ; 1.
DR SMART: SM00225; BTB; 1.
DR PROSITE: PS50097; BTB; 1.
DR PROSITE: PS00036; BZIP_BASIC; 1.
KW Transcription regulation; Activator; Repressor; DNA-binding;
KW Nuclear protein.
FT DOMAIN 37 103 BTB.
FT DNABIND 162 168 POLY-GLU.
FT DNABIND 527 542 BASIC MOTIF.
FT DOMAIN 550 572 LEUCINE-ZIPPER.
SQ SEQUENCE 716 AA; 78935 MW; 913283731AE24333 CRC64;

Query Match 4.0%; Score 94; DB 1; Length 716;
Best Local Similarity 23.1%; Pred. No. 7.2; Matches 101; Conservative 40; Mismatches 140; Indels 156; Gaps 24;

14 SCFAIATAQAKAVNPVAFV---DEV--RSENDLG-----QDNELPIDVQASATQAST 62
124 SCFSFSLQIOLNREGD--LFVCRKDSACORPEDHNSAGEEHEEETMDSTAMACAT 181
63 D-----TANPLDEH-----EPELYTALENKTMILNCALNQ---DIMRLAC 101
182 DQMLDPISFEATAIPIVAKEEALPSESEVPTDKENSEK---GALNQYPRKKYQLAC 237
102 ---YDILVHG-----ETPAVIAKTKRSILN---DET 125
238 TKNVTSADBSHGSGFASFTSESDSPGNSLKPGLPMGOIKSEPSSEETEEESTLCLSGDET 297
126 IWOTIKGRPOVYVOETTPDI-----FLMGNEKGMILTKK 158
298 ---DIKDRPGVDEMRRKQSPARPTSTRTGACADLRSRSSVSPCLRSILFGITKGV---- 350
159 DAKOLEYAKOFTPLSLF-----DIDRNNTPIMWS--RPHNPMVYLPIFMHCK 205
351 ESTGIPSTSOQ--PLVRSACPFNKGISQGDLDKTDYTPLAGNYGQPHVGQKDVSNFAMGS 408
206 PNRSP-----NTPSHARQFTPNFERAPELK-----FGYSVKYKAAEDLMCTDSDL 251
409 PLRGFGPETLCEFSSSPCSGARFLATEHQEPGLMGDGMYNQVPOIK--CEOSYGTNSSD 467
252 WFGYTOQSHWOJFNCKNSRPFVNDYOPEILFLOPV--YSDLPMGDKVMIKMGAVHHNS 309
468 ESG-----SFSRADSSCPVQORGOEVKLPFPVDOITDLPFRNDFQMIMK---HKL 516
310 GESAKL-----SRSMNR 321
517 SEOLEFIHDIRRRSKNR 533

RESULT 11
DP3A_BACSU STANDARD: PRT; 1115 AA.
AC 034623;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE DNA POLYMERASE III ALPHA SUBUNIT (EC 2.7.7.7).
GN DNAME.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/clostridium group;
OC Bacillus/staphylococcus group; Bacillus.

OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98048467; PubMed=9387221;
RX Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;
RT "Sequencing and functional annotation of the Bacillus subtilis genes
in the 200 kb trnB-dnaB region";
RL Microbiology 143:3431-3441(1997).
CC -1- FUNCTION: DNA POLYMERASE III IS A COMPLEX, MULTICHAIN ENZYME
RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN BACTERIA.
CC THIS DNA POLYMERASE ALSO EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
CC THE ALPHA CHAIN IS THE DNA POLYMERASE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE -
CC N PYROPHOSPHATE + DNA(N).
CC -1- SUBUNIT: CONTAINS A CORE (COMPOSED OF ALPHA, EPSILON, AND THETA
CHAINS) THAT ASSOCIATES WITH A TAU SUBUNIT WHICH ALLOW THE CORE
CC DIMERIZATION TO FORM THE POLIIT' COMPLEX. POLIIT' ASSOCIATES WITH
CC THE GAMMA COMPLEX (COMPOSED OF CHAINS GAMMA, DELTA, DELTA', PSI,
CC AND CHI) AND WITH THE BETA CHAIN (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-C FAMILY. DNAME
CC SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL: AF008220; AAC00338.1; -
DR EMBL: Z99118; CAB14883.1; -
DR Subtilist: BG12583; dnaE.
DR InterPro: IPR003141; PHP_N.
DR InterPro: IPR002309; tRNA-synt-2.
DR Pfam: PF02231; PHP_N; 1.
DR Pfam: PF01336; tRNA_ant1; 1.
DR SMART: SM00481; POLIITAC; 1.
KW Transferrase; DNA-directed DNA polymerase; DNA replication;
KW Complete proteome.
SQ SEQUENCE 1115 AA; 125349 MW; E8B04E3398E512FE CRC64;

Query Match 4.0%; Score 94; DB 1; Length 1115;
Best Local Similarity 21.7%; Pred. No. 13; Matches 100; Conservative 62; Mismatches 155; Indels 144; Gaps 26;

20 AIQOAKAVYN-FVAIVDEVREND-----LGQDNELPIDVQASATQASTDTAN- 66
437 ADQLAKLIPSRGWTLDARQOSPOLDKRLRESSLQOYYSIAKKIEGIPRHASTHAGV 496
67 -----PLDEHEPLLYTT--ALENKTMLINCASALNODIMRLACYDVLVHETPAV 113
497 VLSEBPLTYVVLQEGHEBITYLOYAMDH---LEDIGLAKMFLGLRLN-LTLESTT-SM 551
114 IKTKRSIRLDETTIWTIKGPQVYVOETTPDIPLMGNEKGMILTKDAK---OLEYAA--- 167
552 IEKEENIKIDLS-----SISYSD--DKTF-----SLTSKGDITGIFQLESAGMRS 594
168 ---KQFTPLSLSPDLDNRNTPLMSSRPHNPMVYLPIFMHCKPKRSP-NTPSHARQFTPNE 224
595 VLKRLKPSGLE-DIVAUNA--LYRP-GPMENIPLEIDRRKGRAPVHPHEDLRILEDT 649
225 F-----RAPELKFQSVYKAAEDMAGTDSLDLMEFGYTOOSHMQ 262
650 YGVITYQEDIMIASMAGFSLGEADLRLRAVSKKKELD-----RERSHP- 696
263 IFNKNRSPFRVNDYOPEILFLOPVYSDLPMGDKVMIKMGAVHHNSGBSASAKLSRWNR 322
697 -VEGCLAKREYSV-DTANEVY-----DLIYKFANYG-----FNRSNAVA 732
323 YIMAGHEMKNLTVPRIMGRIRKEGSGSQPDNDPILDIYGY-----GDVRFYLQLENKSN 378

DB 733 YSMIGCO---LATLKAHPIYIFMCGLLTSYIGNEDKISQLYTEAKSGCITILPSPVKKSS 789
 QY 379 I-----SGTVRYNPRSGKGLQLDVYPLGKISGYFQIFQ 414
 DB 790 FPFTVNGSVYSLRAIKSV-----GVSAYVDIYK 819
 RESULT 12
 SRK6_BRAOL STANDARD; PRT; 849 AA.
 AC 009092;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE PUTATIVE SERINE/THREONINE KINASE RECEPTOR PRECURSOR (EC 2.7.1.37)
 DE (S-RECEPTOR KINASE) (SRK).
 GN SRK6.
 OS Brassica oleracea (Cauliflower).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Brassica.
 OC NCBI_TaxID=3712;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. S656; TISSUE=stigma;
 RA MEDLINE=92020942; PubMed=1681543;
 RX Stein J.C., Howlett B., Boyes D.C., Nasrallah M.E.;
 RT "Molecular cloning of a putative receptor protein kinase gene encoded
 at the self-incompatibility locus of Brassica oleracea.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:8816-8820(1991).
 CC -I- FUNCTION: INVOLVED IN SPOROPHYTIC SELF-INCOMPATIBILITY SYSTEM
 (THE INABILITY OF FLOWERING PLANTS TO ACHIEVE SELF-
 FERTILIZATION), PROBABLY ACTING IN COMBINATION WITH S-LOCUS-
 SPECIFIC GLYCOPROTEINS. INTERACTION WITH A LIGAND IN THE
 EXTRACELLULAR DOMAIN TRIGGERS THE PROTEIN KINASE ACTIVITY OF THE
 CYTOPLASMIC DOMAIN.
 CC -I- CATALYTIC ACTIVITY: ATP + A PROTEIN - ADP + A PHOSPHOPROTEIN.
 CC -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -I- TISSUE SPECIFICITY: PREDOMINANTLY IN THE PISTIL AND ANTHER.
 CC -I- POLYMORPHISM: THERE ARE A NUMBER OF DIFFERENT S ALLELES IN
 B. OLERACEA, POSSIBLY PROVIDING THE RECOGNITION SPECIFICITY.
 CC -I- SIMILARITY: THE EXTRACELLULAR DOMAIN IS SIMILAR TO S-LOCUS
 GLYCOPROTEINS OF BRASSICA, WHILE THE INTRACELLULAR DOMAIN IS
 A SER/THR-PROTEIN KINASE RELATED TO RAF KINASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M76647; AAA33000.1; ALT_TERM.
 DR HSSP: P11362; IEG1.
 DR InterPro: IPR001480; B_lectin.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR003609; Pan_app.
 DR InterPro: IPR002290; Ser_thr_kin_actsite.
 DR InterPro: IPR000858; Slocus_glycop.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF00954; S_locus_glycop; 1.
 DR SMART: SM00108; B_lectin; 1.
 DR SMART: SM00473; PAN_AP; 1.
 DR SMART: SM00221; STYKC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00011; PROTEIN_KINASE_SF; 1.
 DR TRANSFERASE: Serine/threonine-protein kinase; Signal: ATP-binding;
 KW Transmembrane; Receptor; Glycoprotein; Self-incompatibility.
 FT SIGNAL 1 32 POTENTIAL.
 FT CHAIN 33 849 PUTATIVE SERINE/THREONINE KINASE

FT DOMAIN 33 446 RECEPTOR.
 FT TRANSMEM 447 466 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 467 849 POTENTIAL.
 FT DOMAIN 528 779 CYTOPLASMIC (POTENTIAL).
 FT NP_BIND 534 542 PROTEIN KINASE.
 FT BINDING 556 556 ATP (BY SIMILARITY).
 FT ACT_SITE 653 653 ATP (BY SIMILARITY).
 FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 389 389 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 849 AA; 97231 MW; 7E156059EDDF4370 CRC64;
 Query Match 3.9%; Score 92.5; DB 1; Length 849;
 Best Local Similarity 21.9%; Pred. No. 12; Mismatches 121; Gaps 24;
 Matches 78; Conservative 50; Mismatches 107; Indels 121; Gaps 24;
 QY 105 LVHGETPAVVKTKRSIRLDETIMQTIKQPVVYQETTDPIFLMG---NEK---GMLTK 157
 DB 24 LIHPLALSTIYNT-----LSTFESTLISSNKTLV---SPGSFEVGFPRFTRNRMVGLGWYK 75
 QY 158 K-DAKOLEYAAKQFTPLSLFSD-----LDRNNTPLMSSRPINPMVLPFIEMHGK 205
 DB 76 KVSRTYVWVANRPNPLSNAIGLKISGNMLVLDDHNSKRPWMWN-----LTRGN 125
 QY 206 PNRSP-----NTPSHEARQFT-----PNEPRAPLKFQVSYKVAADLW 245
 DB 126 -ERSPVVAELLANGNEVYWRDSSNDASELYMQSFDPFDLPILPMKILGYNLKT----- 177
 QY 246 GTDSDLMEFGYTOQSHMOIFNGKNSRPFVHDYQ-----PEIFLQO---PYUSDLPMDGK 296
 DB 178 GLNRF-----TSWRSSDDPSGNGF---SKLETSQSLPEFYLSRENFPHRSGPNWG- 226
 QY 297 VRMIGKGAVH-----HSNGESAKLSKSNRAT-----LMAQEMKLYTMP--RI 339
 DB 227 IREFGIDEDOKLSYVYNEFIENNEVAYTFEFTNNSEYSRLLTLLSEGFORTLYWPSIRI 286
 QY 340 WGRIFKFGSGSQPDNDNDIIDY-----YGVGVAFVLELENKSNISGTVR-YNPRS 389
 DB 287 WNRWF-----SSPYD-PQCDTYIMCGPRAYCDV-----NTSPVCNCIGQFNPRN 329
 RESULT 13
 MET_MOUSE
 ID MET_MOUSE STANDARD; PRT; 1379 AA.
 AC P16056; O62125;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HEPATOCYTE GROWTH FACTOR RECEPTOR PRECURSOR (EC 2.7.1.112) (MET PROTO-
 DE ONCOGENE TYROSINE KINASE) (C-MET) (HGF RECEPTOR) (HGF-SR RECEPTOR).
 GN MET.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88262253; PubMed=2838789;
 RA Chan A.M.L., King H.W.S., Deakin E.A., Tempest P.R., Hilkens J.,
 RA Croezen V., Edwards D.R., Wills A.J., Brookes P., Cooper C.S.;
 RT "Characterization of the mouse met proto-oncogene.";
 RL Oncogene 2:593-599(1988).
 [2]
 RP SEQUENCE OF 1199-1270 FROM N.A.
 RX MEDLINE=90152381; PubMed=2482828;
 RA Wilks A.F., Kudban R.R., Hovens C.M., Ralph S.J.;
 RA "The application of the polymerase chain reaction to cloning members

of the protein tyrosine kinase family.";
 RL Gene 85:67-74(1989).
 CC -1- FUNCTION: RECEPTOR FOR HEPATOCYTE GROWTH FACTOR. HAS A TYROSINE-
 CC PROTEIN KINASE ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
 CC TYROSINE PHOSPHATE.
 CC -1- SUBUNIT: HETERODIMER FORMED OF AN ALPHA CHAIN (50 KDA) AND A BETA
 CC CHAIN (145 KDA) WHICH ARE DISULFIDE LINKED.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- DISEASE: ACTIVATION OF MET AFTER REARRANGEMENT WITH THE TPR
 CC (TRANSLOCATED PROMOTER) LOCUS OF CHROMOSOME 1 PRODUCES AN
 CC ONCOGENIC PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE MET TYROSINE KINASE FAMILY OF
 CC RECEPTORS.
 CC -1- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y00671; CA668680.1; -;
 DR EMBL: M33424; AAA40015.1; -;
 DR PIR: S01254; S01254.
 DR HSSP: P11362; 1FG1.
 DR MGD: MGI:96969; Met.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002809; IPT_TTG.
 DR InterPro: IPR003659; PST.
 DR InterPro: IPR002165; Plexin_repeat.
 DR InterPro: IPR001627; Sema.
 DR InterPro: IPR001245; Tyr_kin.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF01437; Plexin_repeat; 1.
 DR Pfam: PF01403; Sema; 1.
 DR Pfam: PF01833; TIG; 3.
 DR PRINTS: PR00109; TYRKINASE.
 DR SMART: SM00429; IPT; 4.
 DR SMART: SM00423; PST; 1.
 DR SMART: SM00219; TYRK; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KM Transferrase; Tyrosine-protein kinase; Proto-oncogene; ATP-binding;
 KM Receptor; Transmembrane; Glycoprotein; Phosphorylation; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 1379
 FT DOMAIN 25 931
 FT TRANSMEM 932 954
 FT DOMAIN 955 1379
 FT DOMAIN 55 499
 FT SITE 1076 1343
 FT NP_BIND 1082 1090
 FT BINDING 1108 1108
 FT ACT_SITE 1202 1202
 FT MOD_RES 1233 1233
 FT CARBOHYD 45
 FT CARBOHYD 106 106
 FT CARBOHYD 201 201
 FT CARBOHYD 357 357
 FT CARBOHYD 398 398
 FT CARBOHYD 404 404
 FT CARBOHYD 606 606
 FT CARBOHYD 634 634
 FT CARBOHYD 784 784
 FT CARBOHYD 878 878
 FT CONFLICT 1199 1199
 FT CONFLICT 1255 1255
 FT CONFLICT 1261 1261

FT CONFLICT 1269 1270 VL -> IP (IN REF. 2).
 SQ SEQUENCE 1379 AA; 153548 MW; FC5CC87FDD8ADEB8 CRC64;
 Query Match 3 9%; Score 92.5; DB 1; Length 1379;
 Best Local Similarity 21.3%; Pred. No. 24;
 Matches 80; Conservative 48; Mismatches 123; Indels 125; Gaps 22;
 QY 39 SENDLGQDDELPI-DVOSATOSASTPTDPADEHEHELETYTTALENTMLNCSALNODIM 97
 DB 1014 SONGACROQYPLTDLSPILTSQSDSDISSPL-----LQN-TVHIDLSALNPELV 1061
 QY 98 RLACDTLVHGETPAVVKTKRSIRLDETITWQTKGKPVQVYQETTDPIFLMGNEKMLTK 157
 DB 1062 QAVOH--VVIGSSSLT-----VHENEVIG--RGHFGCVYH-----GTLTD 1097
 QY 158 KAKOLEYAAKQFTPLSLFDDRNTPIMS-----SRPH-----NPMVY 197
 DB 1098 NDKKTHCAVKSLSNRIT---DIEVSQFLTEGITMKDFHPVLSLGLCLSGSEGPLV 1154
 QY 198 LPIFMHGKPNRSPNPSHARQFTPNFRAPAEIKFQVSVKVAEDLMCTSDPLMFGYTO 257
 DB 1155 LPMYKMG-----DLRNFIRNETHNPYK-----DLIG-----FGL-- 1184
 QY 258 QSHWQIFNGKN---SRPFVHDYQPEIFLTQPYVSDLPMDGKVRMTGKC-----AV 305
 DB 1185 ---QVAKGMKYLASKRFVHRD---LAA RNCLMDEKFTVYKADFGIARDMYKEYYSV 1235
 QY 306 HHSNGESAKLSRSMNRAIYLMAGMEKNLTVMPIWG-----RIFEKSGSDPDNDP-DI 358
 DB 1236 HNKTG--AKLPYKW---MALESLOTQKFTTKSDWSFGVLLMELMTGAPPPDVNTEFI 1290
 QY 359 LDYGGYGVDFLYTOLE 374
 DB 1291 TITLQGG--RRLQPE 1304
 RESULT 14
 YL70_ARCFU
 ID YL70_ARCFU STANDARD; PRT; 324 AA.
 AC 028112;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOTHETICAL PROTEIN AF2170.
 GN AF2170.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
 OC Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Toml J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kervilavege A.R., Graham D.E., Kyriades N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirtness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Sprilys T., Artlich P., Kalne B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Weese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 CC -1- SIMILARITY: CONTAINS 2 KELCH REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:18:34 ; Search time 44.45 Seconds
(without alignments)
757,461 Million cell updates/sec

Title: US-09-787-083-6

Perfect score: 2360

Sequence: 1 MKVSLSTLTLSTLSCFAILA.....YHNEATSPGVGLMDNMGL 442

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR:68:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	815	34.5	382	2	E81195 phospholipase A1,
2	815	34.5	409	2	H81831 probable phospholip
3	373	15.8	286	2	B36971 outer membrane pho
4	366	15.5	329	2	D81279 phospholipase A1 (
5	358	15.2	289	2	A36971 outer membrane pho
6	357	15.1	289	1	PSKCA1 phospholipase A1 (
7	357	15.1	289	2	E86069 outer membrane pho
8	321.5	13.6	289	2	C36971 outer membrane pho
9	246.5	10.4	355	2	C64582 phospholipase A1 P
10	240.5	10.2	355	2	H71930 probable phospholip
11	110.5	4.7	602	1	TVTRRR protein kinase (EC
12	109.5	4.6	800	2	A29003 cellulase (EC 3.2.
13	109.5	4.6	822	2	JT0611 cellulase (EC 3.2.
14	108	4.6	824	3	JC7532 cellulase (EC 3.2.
15	105.5	4.5	901	2	T20122 hypothetical prote
16	103	4.4	783	2	JC5467 cellulase (EC 3.2.
17	101	4.3	719	2	A42893 penicillin-binding
18	101	4.3	4307	2	T20721 hypothetical prote
19	99.5	4.2	1658	2	T42642 phosphotriostide 3
20	99	4.2	719	2	S28031 penicillin-binding
21	99	4.2	5005	2	S28032 penicillin-binding
22	99	4.2	5005	2	F82884 penicillin-binding
23	98.5	4.2	660	2	E83656 methionyl-tRNA syn
24	98.5	4.2	810	2	A49744 AMP deaminase (EC
25	98	4.2	719	2	S28034 penicillin-binding
26	98	4.2	719	2	S28032 penicillin-binding
27	97.5	4.1	1478	2	S20117 protein kinase BCK
28	96.5	4.1	797	2	T46737 X-Pro dipeptidyl-P
29	96.5	4.1	838	2	A96557 probable receptor

30	96	4.1	608	2	S28036 penicillin-binding
31	95	4.0	719	2	S28035 penicillin-binding
32	95	4.0	719	2	S28038 penicillin-binding
33	95	4.0	719	2	S28037 penicillin-binding
34	95	4.0	765	2	T35719 cellulase - Strept
35	94.5	4.0	397	2	A35136 cellulase (EC 3.2.
36	94.5	4.0	601	2	T26062 hypothetical prote
37	94.5	4.0	655	2	T26061 hypothetical prote
38	94	4.0	1115	2	D69617 DNA polymerase III
39	93.5	4.0	564	2	T40777 ferric reductase t
40	93	3.9	791	2	H96839 hypothetical prote
41	92.5	3.9	857	1	A41369 s-receptor kinase
42	92.5	3.9	1379	1	S01254 hepatocyte growth
43	92	3.9	324	2	B69521 hypothetical prote
44	92	3.9	888	2	S50801 AMP deaminase homo
45	91.5	3.9	486	2	S30959 gene 14 protein -

ALIGNMENTS

RESULT	1	ALIGNMENTS
E81195	phospholipase A1, probable NMB0464 [imported] - Neisseria meningitidis (strain MC58 s	
C:Species: Neisseria meningitidis		
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001		
C:Accession: E81195		
R:Telletlin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, Hickey, E.R.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B. rl, H.; Qin, H.; Yamathayan, J.; Gill, J.; Scariato, V.; Maignanl, V.; Pizze, M. Science 287, 1809-1815, 2000		
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58. A:Reference number: A81000; MUID:20175755		
A:Accession: E81195		
A:Status: preliminary		
A:Molecule type: DNA		
A:Residues: 1-582 <TEP>		
A:Cross-references: GB:AE002403; GB:AE002098; NID:g7225688; PID:NAAF40901.1; PID:g722		
A:Experimental source: serogroup B, strain MC58		
C:Genetics:		
A:Gene: NMB0464		
Query Match	34.5%	Score 815; DB 2; Length 382;
Best Local Similarity	44.7%	Pred. NO. 1.9e-58;
Matches 163; Conservative 67; Mismatches 111; Indels 24; Gaps 8;		
OY 87	INCSALNDIMRIACYDVLHGEPPAVT-----KTKRSIRDETIMOTI-RGKPOVYOE 140	
DB 33	LOCALADNVTIRLACYDRIFAAQLPSAGQESKAVLNLLETIRSSLDKGEAVIYVEK 92	
OY 141	TTPDFLNGNKGMLTKDAKOLEYAKQFTPLSLSPDLDRNN--TPLWSSRPNNMYLVP 199	
DB 93	GGDAL-----PADSAGETADITPLSLMYLDDKDKDGLGIVRRHNMVYMLP 139	
OY 200	IFMGKPNRSPNTPSHEAR-OFTPNFRAPAEIKFOVSKVKAEDLWGTDSDLAFGYTQO 258	
DB 140	LMYNNSPVPYAPGSPRTGTVOEKEFGQKRAETKIAEDLFTFRADLMWGYTOR 199	
OY 259	SHWOFN-GKNSRPVRVADYDPEITFLTOPVYSDLPMDCKVRRIGAVHNSGESAKISR 317	
DB 200	SDWQYVNGRKSAPPRNTDYRPEITFLTOPVAKADLPFGRLRLGAFVHQSGQSRPSR 259	
OY 318	SMNRAYLMAAGMEWKNLTVPRIWGRIFKEGSGQFPDDPDLIDYGYGDVFLYOLENKS 377	
DB 260	SMNRITMAAGMEWKGTLITIPVWVARAFDQ-SGDK-NQMDPLADWGYGDVFLYOLNRQ 317	
OY 378	NISGTVRYNPRSSGALQDLYVPLGKISGTFQIGYGGSLDYNHEATSPGVGLMN 437	
DB 318	NVSYLRYNPRPTGAGALDAATFPYKGLKGVNRGFGESLIDYNHKONGICIGLMFN 377	
OY 438	DNMGL 442	

Db 378 DLDGI 382

RESULT 2
H81831

Probable phospholipase A2 [imported] - Neisseria meningitidis (strain 22491 serogroup H81831)
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: H81831
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, N.; Holtroyd, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, N.; Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491.
A:Reference number: A81775; MUID:20222556
A:Accession: H81831
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-409 <PAR>
A:Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85240.1; PID:g738065
A:Experimental source: serogroup A, strain 22491
C:Genetics:
A:Gene: NMA2021

Query Match 34.5%; Score 815; DB 2; Length 409;
Best Local Similarity 44.7%; Pred. No. 2.1e-58;
Matches 163; Conservative 67; Mismatches 111; Indels 24; Gaps 8;

QY 87 INCSAINODIMRLACYDTLVHGETPAVI-----KTKRSIRLDETIGTI-KGKPOVYVGE 140
Db 60 LQCAALTDVNTLACTDRIFPAOLPSSACOGESKAVLLETYSLSLKGAVIVVEK 119
QY 141 TDDPIFLMNEKGMKTKDAKOLEVAAKQTPPLSLFSDDRNN-TPLMSSPNPMVLP 199
Db 120 GGDAL-----PADSAGETADITYTPELSTWYDDDKNDLRGLGRENPNWLP 166
QY 200 ITMHGKPNNSPTPSHEAR-QTPNEFPRAPELKQVSVKKAEDIMGDSLMFGYTQ 258
Db 167 LWNNSPNYAPGSPTRGTVOEKEGQOKRAETKLOVSFKSIEDLEFKTRADLMFGYTOR 226
QY 259 SHMOJFN-GKNSRPRVHYOPEIFLTOPYVSLPMDGKRYMIGMAVHNSGESAKLR 317
Db 227 SMOQITNQRKSAFENNTYKPEIFLTOPYKALPFGGRIMGAGFVHQSNGSPRESR 286
QY 318 SNNRAYLACMEKMLTVMPRIWGRIFKESGSPDPNDPILDYGYGVDFRYQLENKS 377
Db 287 SNNRIYAMGMEKGLTYIPRVVAFDQ-SGDK-NDNPDINDYMGYGVKLDYRLNDRQ 344
QY 378 NISGTVRYNPRSGKALQLDYVYPLGKISGYFOIGYQOSLIDYVHEATSFVGSLMN 437
Db 345 NIVSYLVRYNPKGTGAIEMAYTFPIKGLKGVVRFHGYGESLIDYVHKNQIGIGIMFN 404
QY 438 DMWGL 442
Db 405 DLDGI 409

RESULT 3
B36971
Outer membrane phospholipase A (EC 3.1.1.-) precursor - Klebsiella pneumoniae
C:Species: Klebsiella pneumoniae
C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 18-Jun-1999
C:Accession: B36971; S40129
R:Brok, R.G.P.M.; Brinkman, E.; van Bortel, R.; Bekkers, A.C.A.P.A.; Verheij, H.M.; Tomm, J. Bacteriol. 176, 861-870, 1994
A:Title: Molecular characterization of enterobacterial plid genes encoding outer membrane phospholipase A
A:Reference number: A86971; MUID:94131966
A:Accession: B36971
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-286 <BRO>
A:Cross-references: EMBL:X76901; NID:g436880; PIDN:CAA54223.1; PID:g436881

A:Note: authors translated the codon AAG for residue 112 as Arg
C:Genetics:
A:Gene: plid
C:Superfamily: bacterial phospholipase A1
C:Keywords: carboxylic ester hydrolase

Query Match 15.8%; Score 373; DB 2; Length 286;
Best Local Similarity 42.1%; Pred. No. 9.6e-23;
Matches 90; Conservative 27; Mismatches 87; Indels 10; Gaps 5;

QY 226 RAPELKQVSVKKAEDIMGDSLMFGYTQOSHWQIFNGKSRPRVNDYQPEIFLTQ 285
Db 80 RKDEYKFOLSLAPPLWKGILGDSLSASTQKSMQLSKESAPREFRYEPQLELGF 139
QY 286 PV-YSDLPMDGKRYMIGMAVHNSGESALSSNNRAYLACMEKMLTVMPRIWGRIF 344
Db 140 ATDYQFAGW--TLRDIEMGNHDSNGSDPTSSWNRILYRLAQNQNLVYKPPW---Y 194
QY 345 KEGSGSPDPNDPILDYGYGVDFRYQLENKSNISGTVRYNPRSGKALQLDYVYPLGK 404
Db 195 VVGS---TDDNPDITKYMGYRLKGVQL-GEALISAQGYNNWNTGYGALGVSPITK 250
QY 405 GISGYFOIGYQOSLIDYVHEATSFVGSLMN 438
Db 251 HVRATYQIYSYGESLIDYNFNOTRVGVGLMN 284

RESULT 4
D81279
Phospholipase A1 (EC 3.1.1.32) C11351 [imported] - Campylobacter jejuni (strain NCTC C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 18-Aug-2000
C:Accession: D81279
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chli, C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Ba Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals A:Reference number: A81250; MUID:20150912
A:Accession: D81279
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-329 <PAR>
A:Cross-references: GB:AL139078; GB:AL111168; NID:g6968723; PIDN:CAB73778.1; PID:g6966
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: plidA; C11351
C:Keywords: carboxylic ester hydrolase

Query Match 15.5%; Score 366; DB 2; Length 329;
Best Local Similarity 33.2%; Pred. No. 4.4e-22;
Matches 105; Conservative 48; Mismatches 125; Indels 38; Gaps 13;

QY 130 IKGKPOVYVGETTDPPIFLMNEKGMKTKDAKOLEVAAKQTPPLSLFSDDRNT--PLW 187
Db 43 LKNSVLSISOENNSSQATOTONSITIKKEKQ-----DFSRLANLYGENESEPNLG 96
QY 188 SSRPNHMYVLYPIFMHCKPRSPPTPSHEARQTPNEFPRAPELKQVSVKKAEDIMGT 247
Db 97 IS-SYKNNIFLP-FAVSFNSLGVNNKSEA-----KFOLSVKKRLEFENLGL 141
QY 248 DSDLMFGYTQOSHWQIFNGKSRPRVNDYQPEIFLTOPYV-SDLPMDGKRYMIGMAVH 306
Db 142 DKRYIAYTQTSWMOIY--BHSFPRTYQPEIFLTOPYKALPFGGRIMGAGFVHQSNGSPRESR 196
QY 307 HNSGESAK--LSSNRAYLACMEKMLTVMPRIWGRIFKESGSPDPNDPILDYGY 364
Db 197 ESNKGKDENDQSSRNRIYVSTAILYKFLFVRLMYRI--DENKKDDONPAIILHMGN 253
QY 365 GDYRFLYQLENKSNISGTVRYNPR--SGKALQLDYVYPLGKISGYFOIGYQOSLI 421
Db 254 FDNVLAI-LGDDYFIFNLMRLNNLKFHNNKGAIOVDLGYDIFNNGITWYLDYFNGYGESLI 312


```
QY 295 GAVRMIGAVHNSGES-----AKLSRNRAYLMAGMEKNLTVMPRING 341
DB 411 GTRKQGFV-----NDSDVEDVVIENEGALKLSGLDSNDVSEGNVAMNRLSADGNG 464
QY 342 R 342
DB 465 K 465

RESULT 13
JT0611
cellulase (EC 3.2.1.4), alkaline - Bacillus sp. (strain KSM-64)
N:Alternate names: endo-1,4-beta-glucanase
C:Species: Bacillus sp.
C:Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 10-Dec-1999
C:Accession: JT0611
R:Sumitomo, N.; Ozaki, K.; Kawai, S.; Ito, S.
Biosci. Biotechnol. Biochem. 56, 872-877, 1992
A:Title: Nucleotide sequence of the gene for an alkaline endoglucanase from an alkalophilic
A:Reference number: JT0611; MUID:92305459
A:Accession: JT0611
A:Molecule type: DNA
A:Residues: 1-822 <SUM>
A:Cross-references: GB:M64963; NID:9289264; PIDN:AAA73189.1; PID:9289266
C:Function:
A:Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as
C:Pathway: cellulose degradation
C:Superfamily: Bacillus sp. alkaline cellulase; Thermotoga xylanase A amino-terminal ref
C:Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation
F:583-726/Domain: Thermotoga xylanase A amino-terminal repeat homology <TXA>

Query Match 4.6%; Score 109.5; DB 2; Length 822;
Best Local Similarity 19.8%; Pred. No. 1.2; Mismatches 163; Indels 169; Gaps 25;
Matches 95; Conservative 54;

QY 9 TSLISCFALIAIQAKAVNPVAFVDEVSEND---LCQDNELPIDVQATQASTDRA 65
DB 7 TKQLISSILILVL-LTSLPPALAAAGNTRDNFKHLGNDNVKRRSEAGALQLOEVQDQ 65
QY 66 NPL-DEHEPELY-----TALENKTKMLNCSAL-----NODIMLACY-DLVHGEHP 111
DB 66 MTLVDDHGKRIQLRGKSTHGLQWFPPEILNDNAVKALANDWESNMIRLAMYVGEENGASNP 125
QY 112 AVIKTKRSIRLDETI-----WQT-----IKGKPOVY 138
DB 126 ELIKSVIVIGIDALAIENDMYIVDMVHNAFGPRDPVYAGAEFFRDIAALVPPNNHITY 185
QY 139 QETTPD-----IFLMGNEKGLTKKQAKOLEYAKOFTPLSLFDDLRN---NTPLM 187
DB 186 ELANPESSNNNGAGIIPNNEEGMAVK-----EYADPIVEMLRDSGNADNIIIVGSPWM 240
QY 188 SSRP-----HNPMVLPPIF--MHGKPNRS--PNTPSHEARQFTNEPRAPBLKQ 233
DB 241 SCRPDLADNPIDDHHTMTVHFYTGSHAASTESTPEPTPSNRGVMSNTRFYA---LE 296
QY 234 VSVKVAKAEDLMGT-----DSDLWFGYTOQ-----SHMOIFNGKNS---RP 271
DB 297 NCVAVAFATE--WGTSQANGDGGPYFDEADVWIEFLNENNISMAMWSLTN--KNEVSATFP 353
QY 272 FNVH-----DYQPEIFLTQPVYSDLPW---D 294
DB 354 FELGKSNATSLDPGPQVAVPEELSLSGEYVARIKGVNVEP---IDRTKYTKVLMDPND 410
QY 295 GAVRMIGAVHNSGES-----AKLSRNRAYLMAGMEKNLTVMPRING 341
DB 411 GTRKQGFV-----NDSDVEDVVIENEGALKLSGLDSNDVSEGNVAMNRLSADGNG 464
QY 342 R 342
DB 465 K 465
```

```
RESULT 14
JC7532
cellulase (EC 3.2.1.4), alkaline - Bacillus sp. (strain KSM-5237)
N:Alternate names: endo-1,4-beta-glucanase
C:Species: Bacillus sp. (strain KSM-5237)
C:Date: 31-Mar-2001 #sequence_revision 31-Mar-2001 #text_change 31-Mar-2001
C:Accession: JC7532; PC7107
R:Hakamada, Y.; Hatada, Y.; Koike, K.; Yoshimatsu, T.; Kawai, S.; Kobayashi, T.; Ito,
Biosci. Biotechnol. Biochem. 64, 2281-2289, 2000
A:Title: Deduced amino acid sequence and possible catalytic residues of a thermostable
A:Reference number: JC7532; MUID:21036886
A:Accession: JC7532
A:Molecule type: DNA
A:Residues: 1-824 <HAK>
A:Cross-references: DDBJ:AB018420
A:Experimental source: strain KSM-5237
A:Accession: PC7107
A:Molecule type: protein
A:Residues: 31-50 <HA2>
C:Comment: This enzyme is a thermostable, alkaline enzyme used as effective detergent
C:Genetics:
A:Gene: Egl-237
C:Keywords: hydrolase; glycosidase

Query Match 4.6%; Score 108; DB 3; Length 824;
Best Local Similarity 19.8%; Pred. No. 1.5; Mismatches 159; Indels 166; Gaps 28;
Matches 95; Conservative 60;

QY 9 TSLISCFALIAIQAKAVNPVAFVDEVSEND---LCQDNELPIDVQATQASTDRA 65
DB 7 TKQLISSILILVL-LTSLPPALAAAGNTRDNFKHLGNDNVKRRSEAGALQLOEVQDQ 65
QY 66 NPL-DEHEPELY-----TALENKTKMLNCSAL-----NODIMLACYDLVHGE--- 109
DB 66 MTLVDDHGKRIQLRGKSTHGLQWFPPEILNDNAVKALSNDSNMIRLAMY---VGENCY 121
QY 110 ---TPAVIKTK-----RSIRLDETI---WQT-----IKGK 134
DB 122 ATNPFLIKORVIGIDELAIENDMYIVDMVHNAFGPRDPVYAGAKDFPRETALVPPNP 181
QY 135 QVYVOETTPD-----IFLMGNEKGLTKKQAKOLEYAKOFTPLSLFDDLRN---N 183
DB 182 HIIVELANPESSNNNGAGIIPNNEEGMAVK-----EYADPIVEMLRKSGNADNIIIVG 236
QY 184 TPLWSSRP-----HNPMVLPPIFEMHGKPNRS--PNTPSHEARQFTNEPRAPBLK-- 231
DB 237 SPWMSQRPDLADNPIDDHHTMTVHFYTGSHAASTESTPSE---TPNSEGVMSNT 291
QY 232 ---FQVSVKVAKAEDLMGT-----DSDLWFGYTOQ-----SHMOIFNGKNS-- 269
DB 292 RYALENGVAVFATE--WGTSQASGCGPYFDEADVWIEFLNENNISMAMWSLTN--KNEVS 348
QY 270 ---RPRPV-----HDYQP-EIPLT-----QPV-----YSDLPW--- 293
DB 349 GAFTPPELKSNTATNDPPDPDHVMAPEELSLSGEYVARIKGVNVEPIDRTYTKVLWMP 408
QY 294 -DGKVMIGMGAHVHNSGESAKLSRNRAYLMAGME-----MKNLVYMPIMR 342
DB 409 NDGTKQ--GFGVNSDSPNKEILIAVDNENNTLKVSGLDVSDVSDGNFWMANRLSANGWK 466

RESULT 15
T20122
hypothetical protein F2583.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T20122; T21324
R:McMurray, A.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z19225
A:Accession: T20122
```


A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-901 <W11>
A:Cross-references: EMBL:270750; PIDN:CAA94750.1; GSPDB:GN00023; CESP:F25B3.1
A:Experimental source: clone C50F4
R:Gardner, A.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z19406
A:Accession: T21324
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-901 <W12>
A:Cross-references: EMBL:270752; PIDN:CAA94758.1; GSPDB:GN00023; CESP:F25B3.1
A:Experimental source: clone F25B3
C:Genetics:
A:Gene: CESP:F25B3.1
A:Map position: 5
A:Introns: 25/3; 47/2; 199/1; 242/1; 315/2; 544/3; 600/1; 626/1; 646/1; 666/3; 706/3; 75

Query Match 4.5%; Score 105.5; DB 2; Length 901;
Best Local Similarity 19.6%; Pred. No. 2.8;
Matches 74; Conservative 60; Mismatches 174; Indels 69; Gaps 15;

QY 31 VAFVDEVRSENDLQ-----DNELPIDVQSATOSASTDTANPLDEHEPELYTTALENK 83
DB 468 VAMITEIRNQKLEAEVDRHPIPEVVPYPPQLASRNPALNOPTDADDDAETSNMRPERS 527
QY 84 TWLI-----NCSALNODIMRLACYDVLHGEPYAVIKTKRSIRLDETIMQT 129
DB 528 NVSITMVTGVAIRASNSASPSSKRDLEQRARDLIEKSTTPAATPNSRKASDEERRREE 587
QY 130 I-----KGPQVYVYQETTDPIFLMGNEKMLTKKDAKQLEYAAKQF-----TPLSLSPD 178
DB 588 VRRMLNEKQHPAIPSTSSSPYPTFRIRIDGSNT--DLRRIELDVHKKRDPSPPLVRKQ 645
QY 179 LDRNNTPLMSSRPHNMYLPIFMHGKPNRSPY--TPSHEA--RQFTPNEFRAPELKF 232
DB 646 YDPNDI-----PH-----VPAIGRGRTGNRRDSSSTPSSASTFEDRVKRYGSMRSALKE 694
QY 233 QVSVKYK---AAEDLMGTSDLMFGYTOQ--SHMQ---IFNGKNSRPFRVHDYQPEIF 282
DB 695 SIQLMAKQGYMGNDPESSSQDALATPTKFSQWMEKDVDDVEGTANELVRIIDERISDIT 754
QY 283 LTPQPVYSDLPMQGVKVMIGMAVHSHNGESAKLSRSMNRAVYLMAGMEWKN-LTVPPIWG 341
DB 755 AOADVYIOD-----KIRETEVGSSEEMLTASYLETNERNTLVHROEYNNIETIRQVTS 809
QY 342 RIFKES--GSQPDNDP 356
DB 810 EIDQLGKQINEVDDFP 826

Search completed: November 30, 2001, 14:18:35
JOB time: 257 sec

This Page Blank (uspto)

QY 89 ----CSALNODIMRLACDYLHGETP---AVIKTRSRIRLDETIMOTIKGRPOVYQE 140
| | | | |
| | | | |
Db 622 VRKICSAID-----GVETPSVTEAVKRLKRAVLP- 652
| | | | |
QY 141 TTDFIFLNGNEKMLTKKAKOLEYAAKQFTPLSLSPDLDRNNPLMSSRHPNMYLP- 199
| | | | |
| | | | |
Db 653 -----NKSADVT-----SLGSDTRKNTSGSLNENPVQVSM 686
| | | | |
| | | | |
QY 200 -----IFMNGKPNR-----SPNTPSHEAKQFTPNFRAPELKFQVSVYKAAEDL- 244
| | | | |
| | | | |
Db 687 HLTTRIYDLRLHANSSRCSTGCPGRSRIKEAWTATE-----QLQFTVYAAHGSS 738
| | | | |
QY 245 -WGTDSPLMFGYTOOSHWOIFNGKN-SRPF--RVHDYQPEIFLTQ-----PV-YSDL 291
| | | | |
| | | | |
Db 739 NWVSNYEKYYILCSLSH-----NGKDLFKPIOSKVGTYKKNFFYLKWDLLIIFPIQISOL 794
| | | | |
| | | | |
QY 292 PMDGKVRMIGGAVHNSGESAKLSRSWN-----RAYLMAEMKKNLTVM 336
| | | | |
| | | | |
Db 795 PLESVHLTLFGLVNLQSSGSSPDSSNKQKGPALGKVSLLTFDKRFLTCG-----TKL 848
| | | | |
QY 337 PRMGRIKESGSGOPDPNDILDYGYGDVRFPLYOLENKNISGTVAYNRSKGALQL 396
| | | | |
| | | | |
Db 849 LYLW-----TSSHTNSIPEAIKPKSVYMERIVLQVDFPSPADIIYTSPOIDRNIIQ 901
| | | | |
| | | | |
QY 397 DYVYPLGKISGYQIFQGYGOSLIDYNHEATSGV 432
| | | | |
| | | | |
Db 902 DKLETLESIDIG-----KLIDIHRDSSFGL 927
| | | | |
| | | | |
RESULT 2
US-09-170-996-30
; Sequence 30, Application US/09170996
; Patent No. 6291220
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Moliz, Lisa
; APPLICANT: Chen, Yen-Wen
; TITLE OF INVENTION: No. 6291220e1 PI 3-Kinase Polypeptides
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/170,996
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/609,049
; FILING DATE: 29-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-063700US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1726 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-170-996-30

Query Match 4.3%; Score 100.5; DB 4; Length 1726;
Best Local Similarity 18.9%; Pred. No. 0.62;
Matches 86; Conservative 66; Mismatches 145; Indels 159; Gaps 21;
QY 45 QDNELPIDV-----QSATQASASTDANPLDE-----HEPLTYTALENKTMLIN----- 88
| | | | |
| | | | |
Db 563 EDDAPVLDLNLKYLOIEKPYKEVMTRHPELIDSYHQVEL-ALQENOHRAVDYIKA 621
| | | | |
| | | | |
QY 89 ----CSALNODIMRLACDYLHGETP---AVIKTRSRIRLDETIMOTIKGRPOVYQE 140
| | | | |
| | | | |
Db 622 VRKICSAID-----GVETPSVTEAVKRLKRAVLP- 652
| | | | |
QY 141 TTDFIFLNGNEKMLTKKAKOLEYAAKQFTPLSLSPDLDRNNPLMSSRHPNMYLP- 199
| | | | |
| | | | |
Db 653 -----NKSADVT-----SLGSDTRKNTSGSLNENPVQVSM 686
| | | | |
| | | | |
QY 200 -----IFMNGKPNR-----SPNTPSHEAKQFTPNFRAPELKFQVSVYKAAEDL- 244
| | | | |
| | | | |
Db 687 HLTTRIYDLRLHANSSRCSTGCPGRSRIKEAWTATE-----QLQFTVYAAHGSS 738
| | | | |
QY 245 -WGTDSPLMFGYTOOSHWOIFNGKN-SRPF--RVHDYQPEIFLTQ-----PV-YSDL 291
| | | | |
| | | | |
Db 739 NWVSNYEKYYILCSLSH-----NGKDLFKPIOSKVGTYKKNFFYLKWDLLIIFPIQISOL 794
| | | | |
| | | | |
QY 292 PMDGKVRMIGGAVHNSGESAKLSRSWN-----RAYLMAEMKKNLTVM 336
| | | | |
| | | | |
Db 795 PLESVHLTLFGLVNLQSSGSSPDSSNKQKGPALGKVSLLTFDKRFLTCG-----TKL 848
| | | | |
QY 337 PRMGRIKESGSGOPDPNDILDYGYGDVRFPLYOLENKNISGTVAYNRSKGALQL 396
| | | | |
| | | | |
Db 849 LYLW-----TSSHTNSIPEAIKPKSVYMERIVLQVDFPSPADIIYTSPOIDRNIIQ 901
| | | | |
| | | | |
QY 397 DYVYPLGKISGYQIFQGYGOSLIDYNHEATSGV 432
| | | | |
| | | | |
Db 902 DKLETLESIDIG-----KLIDIHRDSSFGL 927
| | | | |
| | | | |
RESULT 3
US-08-609-049A-13
; Sequence 13, Application US/08609049A
; Patent No. 5948664
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Moliz, Lisa
; APPLICANT: Chen, Yen-Wen
; TITLE OF INVENTION: No. 5948664e1 PI 3-Kinase Polypeptides
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,049A
; FILING DATE: 29-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-063700US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:

LENGTH: 1658 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-609-049A-13

Query Match 4.2%; Score 99.5; DB 2; Length 1658;
 Best Local Similarity 18.9%; Pred. No. 0.74;
 Matches 101; Conservative 79; Mismatches 194; Indels 159; Gaps 23;

QY 5 LSTLTSLIPCFALIAIO--QAQAVNPVAFVDEVS-----ENDLGQNE 48
 DB 381 LSPVTVQRMNCGENASVKSIEIGLQLPVTFCTDVSTVEIIMQALCWVHDLNO-- 437
 QY 49 LPIDVOS-----ATQASSTDPANPLDEHE--PELYTTALEKMTMLNCSALNDIMRLAC 101
 DB 438 --VDVGSYILKVGQGEVYLNHCLGSHHEIQNCRKMDTEIKQLLTLISMQCLARTAE 495
 QY 102 YDPLVHGEPVAVIKTKRSIRLDETIMQTIKGPQVYQETTPIF----- 146
 DB 496 DD-----EAP-----VDLNKYLQIEKPYKEVMI RHPEELLD SYHYQVELALQTEN 542
 QY 147 -----LMGNEKGM LTKDAKOLEYAA-----KQTPPLSLSDLDNRN 183
 DB 543 QHRAVDQVIAKVRKICSAIDGVEPVSYTEA-VKIKRAVNLPRNKSADVTSLSGSDTRKN 601
 QY 184 TPLMSSRPHPMYLVP-----IFMHGKPNR-----SPNTPSHKAKOPTPNEFRAP 229
 DB 602 STKGSILNPEPNVQVSMDLTAITAYDLRLHANSRSTGCPGRSRNKKEWTATE----- 656
 QY 230 LKQVSVKKAABDL--WGTDSDLWFGYTOQSHMOIFNGKN-SRPF--RVHDYQPEIF 282
 DB 657 --QLOFTVYAAHGISSNWNYSNEYKYLICLSLH--NGKDLFKPIQSKKVGYYKNFEY 709
 QY 283 LTO-----PV-YSDLPMGKVRMIGMGAHVHNSGESAKLSRSWN----- 320
 DB 710 LIKWDELIFPIQISQPLESVLHLTLFGLVNLQSSGSSPDSNKKORKEPALGKVS LTFD 769
 QY 321 -RAYLAGMEKMTLVMPRIWGRIFKESGSGOPDNDIIDDYGYGVDFRFLYOLENKSNI 379
 DB 770 FKRFITCG-----TKLLYLM-----TSSHTNSIPGALPKSYVNERIVLQVDFPSPA 816
 QY 380 SGTVRNPNSGKALQLDYVPLGKGISGYFOIRQGYGOSLIDYNHETATSGV 432
 DB 817 FDIITYSPQIDRNIIQDKLETLESIDIK-----KILDIIHNDSSFGL 859

RESULT 4
 US-09-170-996-13
 Sequence 13, Application US/09170996
 Patent No. 6291220
 GENERAL INFORMATION:
 APPLICANT: Williams, Lewis T.
 APPLICANT: Moliz, Lisa
 APPLICANT: Chen, Yen-Wen
 TITLE OF INVENTION: No. 6291220e1 PI 3-kinase Polypeptides
 NUMBER OF SEQUENCES: 32
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend and Crew LLP
 STREET: Two Embarcadero Center, 8th Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94111-3834
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/170,996

FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/609,049
 FILING DATE: 29-FEB-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Dow, Karen B.
 REGISTRATION NUMBER: 29,684
 REFERENCE/DOCKET NUMBER: 2307K-063700US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-326-2400
 TELEFAX: 415-326-2422
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1658 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-170-996-13

Query Match 4.2%; Score 99.5; DB 4; Length 1658;
 Best Local Similarity 18.9%; Pred. No. 0.74;
 Matches 101; Conservative 79; Mismatches 194; Indels 159; Gaps 23;

QY 5 LSTLTSLIPCFALIAIO--QAQAVNPVAFVDEVS-----ENDLGQNE 48
 DB 381 LSPVTVQRMNCGENASVKSIEIGLQLPVTFCTDVSTVEIIMQALCWVHDLNO-- 437
 QY 49 LPIDVOS-----ATQASSTDPANPLDEHE--PELYTTALEKMTMLNCSALNDIMRLAC 101
 DB 438 --VDVGSYILKVGQGEVYLNHCLGSHHEIQNCRKMDTEIKQLLTLISMQCLARTAE 495
 QY 102 YDPLVHGEPVAVIKTKRSIRLDETIMQTIKGPQVYQETTPIF----- 146
 DB 496 DD-----EAP-----VDLNKYLQIEKPYKEVMI RHPEELLD SYHYQVELALQTEN 542
 QY 147 -----LMGNEKGM LTKDAKOLEYAA-----KQTPPLSLSDLDNRN 183
 DB 543 QHRAVDQVIAKVRKICSAIDGVEPVSYTEA-VKIKRAVNLPRNKSADVTSLSGSDTRKN 601
 QY 184 TPLMSSRPHPMYLVP-----IFMHGKPNR-----SPNTPSHKAKOPTPNEFRAP 229
 DB 602 STKGSILNPEPNVQVSMDLTAITAYDLRLHANSRSTGCPGRSRNKKEWTATE----- 656
 QY 230 LKQVSVKKAABDL--WGTDSDLWFGYTOQSHMOIFNGKN-SRPF--RVHDYQPEIF 282
 DB 657 --QLOFTVYAAHGISSNWNYSNEYKYLICLSLH--NGKDLFKPIQSKKVGYYKNFEY 709
 QY 283 LTO-----PV-YSDLPMGKVRMIGMGAHVHNSGESAKLSRSWN----- 320
 DB 710 LIKWDELIFPIQISQPLESVLHLTLFGLVNLQSSGSSPDSNKKORKEPALGKVS LTFD 769
 QY 321 -RAYLAGMEKMTLVMPRIWGRIFKESGSGOPDNDIIDDYGYGVDFRFLYOLENKSNI 379
 DB 770 FKRFITCG-----TKLLYLM-----TSSHTNSIPGALPKSYVNERIVLQVDFPSPA 816
 QY 380 SGTVRNPNSGKALQLDYVPLGKGISGYFOIRQGYGOSLIDYNHETATSGV 432
 DB 817 FDIITYSPQIDRNIIQDKLETLESIDIK-----KILDIIHNDSSFGL 859

RESULT 5
 US-08-481-435-6
 Sequence 6, Application US/08481435
 Patent No. 6027906
 GENERAL INFORMATION:
 APPLICANT: Balganes, Tanjore S
 APPLICANT: Town, Christine
 TITLE OF INVENTION: No. 6027906e1 Polypeptides
 NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:

ADDRESSER: White & Case
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,435
FILING DATE: 10-JUL-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: IN 580/MAS/94
FILING DATE: 01-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9404072-2
FILING DATE: 24-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Steiner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-151
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 682 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-481-435-6

Query Match 4.08; Score 95.5; DB 3; Length 682;
Best Local Similarity 19.98; Pred. No. 0.45;
Matches 89; Conservative 69; Mismatches 186; Indels 103; Gaps 24;

QY 18 ILATQQAQVNPVAFVDEVRSENDL-GQDNEPLIDVQASQ-----ASTDPANP 67
DB 10 LVATTSKTIYDNKQIADLSESRVNAQANDIPTDLKALVSIEDHAFEDHRCIDITIRI 69
QY 68 LDEHEPELYTALLENKTMILNCSALNODIMRLACYDTLVHGETPAVVIKRSIRLDETIW 127
DB 70 LGARLRNLSQNSISLGG-----SALTQOLIKLYEFSTSTSDQITIS-----RKAQEW 115
QY 128 QTIKGRQVYQET-----TDPIFLMGNEKMLTK-----KDAKOLF-----YA 166
DB 116 LAIDLEKATKQEIILTYIYNKYNVMSNGYMGTAQANYGKDLNNLSLPOLALLAGMPQA 175
QY 167 AKOTPTLS-LSFDDLRRNTPLMSSRPHPVYVLPPIFMGKPNRSPNPSHEAKOFTPEEF 225
DB 176 PNOYDPYSHPEAADRRRLVL--SEMKNOGTI-----SAEQIEKAVNTITGLQSLKAS 229
QY 226 RAPE-----LKFQVS-YVKAEDLMGTDSDLMFGYTQ--QSH-WQIFNGKNSRPFVHD 276
DB 230 NRPAYMONYKLEYINQVBEETGYNLLTGMVYTNVQDEAKHLMDINDTEYVAYPDE 289
QY 277 YQPEIFLTQVYSDLPMDGKVRMTGMCAVHNSN-----GSAKLSRWNNRAYILMAGHEW 330
DB 290 LQ-----VASTIVYVS-NGKV-IAQLGARHQSNSVSGINQAVETNRDM-----GSTM 335
QY 331 KNLV-VMPRMIRGRIFKEGSGSQPDNDPILDYG-----YGDVRFLYOLEKNS 377
DB 336 KPIIDYALALEGYV-ESTATVHDEP--YINPGTNTVYVMMDRGYFSGNITLQIALQQR 392
QY 378 NISGTVRYN-----PRSGKALDLDY 398
DB 393 NVPVETLTKVGLNRAKTFLLNGLIGIDY 419

RESULT 6
US-07-717-331F-2
Sequence 2, Application US/07717331F
Patent No. 5484905
GENERAL INFORMATION:
APPLICANT: June Nasrallah, Michael Nasrallah, and Joshua
APPLICANT: Stein
TITLE OF INVENTION: A Receptor Protein Kinase Gene
TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: Yahwak & Associates
STREET: 25 Skytop Drive
CITY: Trumbull
STATE: Connecticut
COUNTRY: USA
ZIP: 06611
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/717,331F
FILING DATE: June 19th 1991
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: George M. Yahwak
REGISTRATION NUMBER: 26,824
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203)268-1951
TELEFAX: (203)268-1951
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 857 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-717-331F-2

Query Match 3.88; Score 90.5; DB 1; Length 857;
Best Local Similarity 21.44; Pred. No. 2.2;
Matches 76; Conservative 50; Mismatches 110; Indels 119; Gaps 23;

QY 105 LVHGETPAVVIKTKRSIRLDETIWOTIKGPQVYVQETTDPIFLMG---NEK---GMLTK 157
DB 24 LIHPALSIYINT-----LSTFESLTISSNKTLV---SGSIFEVQGFRTNSRWLYGMYWK 75
QY 158 K-DAKQLEYAAKQFTPLSLFD-----LDRNTPMLSSRPHPVYVLPPIFMGK 205
DB 76 KYSDRTYVAVANRDNPLNAGITLKISGNNLVLLDHSKPKVWMTN-----LTRGN 125
QY 206 PNRSP-----NTPSHEAKOFT-----PNEFRAPELKFOVSVKKAEDLM 245
DB 126 -ERSPVVAELLANGFVWRDSSNDASEYLMQSFYPTDITLPEKKGYNLKT----- 177
QY 246 GTDSDLMFGYTQOQSHQWQIFNGKNSRPFVHDYQ-----PELFIHQ---PYVSDLPMDGK 296
DB 178 GINRFL-----TSMRSSDDPSSGNF---SYKLETOQLPEYISRENFPMHRSQPMWG- 226
QY 297 VRMIGMGAVH-----HSNGESAKLSRWNNRAY-----LAAGMKKNTLVMP--RI 339
DB 227 IRFSGIPEDOKLSYVWYVNFIEENNEVAVTFRTNNSFYSRLTLISEGYFORLTWVPSIRI 286
QY 340 WGRIFKEGSGSQPD---NPDILDYGYGQVRFYOLEKNSNISGTVR-YNPRS 389
DB 287 WNRFWSSPVDRQCDIYIMGCP-----YAYCDV-----NTPSPVNCIÖGFNPRN 329

RESULT 7
US-09-091-117-5

Db 213 ---DAFFGSLTATATATPMSAHTKQNERQWNPINSPVLLQTT----- 258
QY 134 MLTKDAKOLEYAAKQTPPLSLFPLDRNNPL---WSSRPNPMYVLPITMGKPNRS 209
Db 259 -----RMT-----DIDRSTSAHGENMTQEPPE----- 283
QY 210 PNTPEAKQTPNFRAPELKFQYSVKYKAEDLM-----GTDSDLMFGTQOS 259
Db 284 -NNHEYODEETPH-----ATSTWADPNSTTEEAATQKEWF-----EN 322
QY 260 HMOIFNGKN-----SRPRVHDYOPEIFLTPVYSDLPM---DGKVRMIGMG 303
Db 323 EMQ---GKNPTPSDSHTBECTASAHNNHPSQKMTTQSOEDVGTDEFDPISHMGGG 379
QY 304 AVHHSNGESA 313
Db 380 HOTESKGHS 389

RESULT 11
US-08-961-083-2
; Sequence 2, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8512
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 666 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-961-083-2

Query Match 3.7%; Score 87.5; DB 4; Length 666;
Best Local Similarity 20.4%; Pred. No. 3;
Matches 87; Conservative 60; Mismatches 176; Indels 103; Gaps 24;

QY 42 DLG-----QDNELPIDVQSATOS-----ASTDTANPLDHEPELYTALENKTM 85
Db 12 DLGSRRAVNAQADIPDLVKAIVSIEDHREFDHRGIDTIRITAGAFRLRLQSNLSLGG-- 69
QY 86 LINCALNDIMRLACYDTLVHGERPAVVKTKRSIRLDETINQTIKGRQVYVQET----- 141

Db 70 ----SITLQOILKLTFFSTSDQITIS-----RKADENALIALDEKAKIKQELLTVY 117
QY 142 TDPIFLMGNEKMLTK-----KDAKQLE-----YAAKQTPPLS-LSFDLDRNN 183
Db 118 INKVVYMSNGNYCMTQAQNYKDKLNNLSLPOLALLAGMPOANQYDPYSHPEAADRNN 177
QY 184 TPLMSRRPNPMYVLPITFMHGKPNRSPNPSHAKQFTNERRAP-----LKFOVS-VK 237
Db 178 LVL--SEMRNQGTY-----SAEQYERKAVNPTITDGLQSLKSASNYPMYMDNYLKEVINQVE 231
QY 238 VAAEEDITDSDLMFGYTO--QSH-WQIFNGKNSRPRVHDYOPEIFLTPVYSDLPM 294
Db 232 EETGYNLITGGADYTYTNVQEAQKHLMDLYNDEYAYPPDDELQ-----VASTYDVS-N 285
QY 295 GKYRMIGMGAVHHSN-----GESAKLSRWBRAYLMAEMKMLT-VMPRI-----W 340
Db 286 GRV-IAQLGARHQSSNVSGINQAVETNRDW-----GSTMKPLTDVAPALEYGVYDST 337
QY 341 GRIFKESGSGOPDNDPILDY--YGY-GDYRFLXQLENKSNISGTVRYN-----PRSGKG 392
Db 338 ATIVHDEPYNPGTNPVYVNMWRGYGNTLQYALQOSRNVPAVETLNKVGILNRAKTFPLN 397
QY 393 ATQLDY 398
Db 398 GLGIDY 403

RESULT 12
US-08-373-134D-2
; Sequence 2, Application US/08373134D
; Patent No. 5780296
; GENERAL INFORMATION:
; APPLICANT: Kmiec, Eric
; APPLICANT: Holloman, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS TO PROMOTE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/373,134D
; FILING DATE: January 17, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Friebe, Thomas F.
; REGISTRATION NUMBER: 29,258
; REFERENCE/DOCKET NUMBER: 7991-007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 781 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-373-134D-2

Query Match 3.6%; Score 86; DB 1; Length 781;
Best Local Similarity 20.7%; Pred. No. 5.5;

Qy	107	HGEPLAVKTKRSLRIDELTIMQITKGRPOVYUOE-TTDP1----	FLMGNEKSGMLTKKDAQ1	162
Db	614	FTQAPSTVPKAT-----	QITSTPEVLQSPSTPEVPFTRTLGAEBEITQTPBSA--	662
Qy	163	LEYAKOFTPLTSLFDDLRNRTPLMSSRP-----	HN-----PMVLPL-----IFMHGK	205
Db	663	---APEYVTRBSSIMTPEAOSTPLASONPFISSGICGINTPRITYPVOTPHHTOKLYTEHK	719	
Qy	206	PNRSPNTPS--HE---AKOFTP--NEFRAPELKOVSVYKVA-----	AEDIMGTD	248
Db	720	TLSPPTVYSEEHKEMSTAESQRPDLDAIVKEKFNDEVEYATACVSYKSPRYVETLNKKVD	779	
Qy	249	-----SDLMFGYUQOOSHQJFNGKNS-----	RPF-----	272
Db	780	LVYVMEDEISGNSPAGVFNSENKMK-----	OKLYYRVTDGRTSVQJMLCLSTSHSPEDYCLFD	836
Qy	273	-----RHADYDPELFLT---QPVYSDLPMDGKV-----	RHIGKGAVHNHN	309
Db	837	TSLIAREKIDAPELYFTSDPOTACITTLPSGVVPREFEWSLNNVSLPEYLTATVVSHTA	896	
Qy	310	GE-----SAKLSRSM	319	
Db	897	GOSTVWKKSSARAGEAW	912	

```

RESULT 15
PCT-US96-03916-66
; Sequence 66, Application PC/TUS9603916
; GENERAL INFORMATION:

```

APPLICANT: Cochran, Mark D.
 TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
 TITLE OF INVENTION: AND USES THEREOF
 NUMBER OF SEQUENCES: 72
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Cooper & Dunham LLP
 STREET: 1185 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US96/03916
 FILING DATE: 23-MAR-1995
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/126,597
 FILING DATE: 24-SEP-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: White, John P.
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 39116-A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 278-0400
 TELEFAX: (212) 391-0525
 INFORMATION FOR SEQ ID NO: 66:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 985 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein

Query Match	3.68;	Score 86;	DB 5;	Length 985;
Best Local Similarity	19.48;	Pred. No. 8;		
Matches	73;	Conservative	47;	Mismatches 126;
				Indels 130;
				Gaps 20

Oy	48	ELPIDVOSATQSDPRDANPLDHEBPELYT-TALENTMILINCALNQDMLACVYTLV	106
Db	563	QTPSTPEKTYTLSSST-----EPALFTTOSAGTEATLTQTSASAPDTRKQSTETHF	613
Oy	107	HGETPAVYKTKRSIRLDETTIMOTIKGPROVYOE-TTDP1---FLIMENEGMLTKKDAQO	162
Db	614	FTQAPSTYPRKAT-----QFTSTEBEVLTGOSPRTEPVPFTRLCAABEITQTPSA--	662
Oy	163	LEYAKOFTPLSLSEDDLRNTPMLMSSRP-----HN-----PMYLP----IFNHKG	205
Db	663	--APEVYTRSSSTMPETAOSTPLASONPTSSGCTGHNTPEPRYPVOTPTHTOKLYTEHK	719
Oy	206	PNRSPNTPS--HE--AKQFPR-NEFARPELKPQVSUKVA-----AEDLIGTD	248
Db	720	TLSFPTVVSSEHEMSTAESQTPRLDVAIVEXKESNDGEVATATCVSYVKSPPYREVTNKKVD	779
Oy	249	-----SDLMEGYUQOSHMOJFNKNS-----RPF-----	272
Db	780	LVDVWDELSGNSPACVFNFSNKM---QKOLYKRTDGRITSQJMLCSTSHSPRPYLF	836
Oy	273	-----RVHNDYQPELFLT--QPVYSDLPMWDGKV-----RMIGMCAVHNHN	309
Db	837	TSLLAREKDIDPELYFTSDPOTACTITLPGGVPRFEMSLNNYSLPEYLTATVVSHTA	896
Oy	310	GE-----SAKLSRW	319
Db	897	GOSTVWKSSARAGEAW	912

Search completed: November 30, 2001, 14:17:51
Job time: 228 sec

2 1 2

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:18:35 ; Search time 44.45 Seconds

(without alignments)
757.461 Million cell updates/sec

Title: US-09-787-083-8

Perfect score: 2363

Sequence: 1 MKVSLSTLTSTILPCFAILA.....YNHEATSEFGVGLMNDMNGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR.68:*

- 1: pir1.*
- 2: pir2.*
- 3: pir3.*
- 4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	815	34.5	382	2	E81195 phospholipase A1, probable phospholip
2	815	34.5	409	2	H81831 outer membrane pho
3	373	15.8	286	2	B36971 phospholipase A1 (
4	366	15.5	329	2	D81279 outer membrane pho
5	358	15.2	289	2	A36971 phospholipase A1 (
6	357	15.1	289	1	P58CA1 phospholipase A1 (
7	357	15.1	289	2	E86069 outer membrane pho
8	321.5	13.6	289	2	C36971 phospholipase A1 P
9	246.5	10.4	355	2	C64582 phospholipase A1 P
10	240.5	10.2	355	2	H71930 probable phosphol
11	111.5	4.7	602	1	TYVTRR protein kinase (EC
12	108.5	4.6	800	2	A28003 cellulase (EC 3.2.
13	108.5	4.6	822	2	JT0611 cellulase (EC 3.2.
14	107.5	4.5	824	3	JC7532 cellulase (EC 3.2.
15	106.5	4.5	901	2	T20122 hypothetical prote
16	100.5	4.3	660	2	E83656 methionyl-tRNA syn
17	100.5	4.3	783	2	JC5467 cellulase (EC 3.2.
18	100.5	4.3	1658	2	T42642 phospholipase A1 (
19	99.5	4.2	797	2	T46737 X-Pro dipeptidyl-p
20	99	4.2	5005	2	F82884 hypothetical prote
21	98.5	4.2	810	2	S49744 AMP deaminase (EC
22	98.5	4.2	838	2	A96557 AMP deaminase (EC
23	96	4.1	791	2	H66839 hypothetical prote
24	96	4.1	4307	2	T20721 hypothetical prote
25	95.5	4.0	719	2	A42893 penicillin-binding
26	95	4.0	765	2	T35719 chitinase - Strept
27	95	4.0	888	2	S50801 AMP deaminase homo
28	94.5	4.0	397	2	A35136 cellulase (EC 3.2.
29	94.5	4.0	601	2	T26062 hypothetical prote

30	94.5	4.0	655	2	T26061 hypothetical prote
31	94.5	4.0	1310	2	T40135 probable involtome
32	94.5	4.0	1478	2	S20117 protein kinase BCK
33	93.5	4.0	564	2	T40777 ferlic reductase t
34	93.5	4.0	719	2	S28031 penicillin-binding
35	93.5	4.0	719	2	S28033 penicillin-binding
36	93.5	4.0	857	1	A41369 S-receptor kinase
37	93	3.9	1078	2	T18352 protein P120 - Msc
38	92.5	3.9	486	2	S30959 gene 14 protein -
39	92.5	3.9	719	2	S28034 penicillin-binding
40	92.5	3.9	719	2	S28032 penicillin-binding
41	92	3.9	324	2	B69521 hypothetical prote
42	92	3.9	470	2	T43675 cog-2 protein - Ca
43	91.5	3.9	796	2	JC7355 peroxisome prolif
44	91.5	3.9	1641	2	D82704 conserved hypothet
45	91	3.9	367	2	T24058 hypothetical prote

ALIGNMENTS

RESULT 1	phospholipase A1, probable NMB0464 [imported] - Neisseria meningitidis (strain MC58 s
E81195	C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000	#sequence: Revision 31-Mar-2000 #text: change 19-Jan-2001
C:Accession: E81195	R:Retelien, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, R.E.; Eisen, Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B. ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignanl, V.; Pizzaz, M. Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.	A:Reference number: A81000; MUID:20175755
A:Accession: E81195	A:Status: preliminary
A:Molecule type: DNA	A:Residues: 1-382 <TEXT>
A:Cross-references: GB:AE002403; GB:AE002098; NID:97225688; PIDN:AAF40901.1; PID:9722	A:Experimental source: serogroup B, strain MC58
C:Genetics:	A:Gene: NMB0464
Query Match	34.5%; Score 815; DB 2; Length 382;
Best Local Similarity	44.7%; Pred. No. 3.5e-58;
Matches 163; Conservative	67; Mismatches 111; Indels 24; Gaps 8;
OY	87 INCSALNODIMRLACVDTLVHGETPANI-----KTKRSITLDETITMOTI-KGRPOVYOE 140
DB	33 LQCAALTDNVTYRLACYDRIFAAQLPSSAGOGESKAVLNLTEVRSISLDKGEAVIYVEK 92
OY	141 TTPPIFLMGNEKMLTKKDAKOLEYAAKOPTPLISFDLRNN-TPLMSRPHNPVLP 199
DB	93 GGNAL-----PADSAGETRADITPTLSIMLDLKDNDRLGLCYREHNPVLP 139
OY	200 IFNHGKPNRSPNTPSH-EAKQFTPNFEFRELKFOYSVKAEDLMGTDSDLMFGYTOQ 258
DB	140 LWNNSPNVAPGSPTRTQVQEFKGOOKRAETKLOVSFKIAEDLFTKTRADLMFGYTOR 199
OY	259 SHMOQIN-GKNSRPFVHDYQPELITQPYVSDLPMDGKVRMIGMCAVHNSNESAKLSR 317
DB	200 SDMOIYNQGRKSAFRTDVKPEIFLQPKADLPFGGRRLMAGAGVHNSQOSRPSR 259
OY	318 SNRRATLMAGMEKNTLVMPRIWGRITFEKSGSQOPDNDPILDYGYGVDFRYOLENKS 377
DB	260 SNRRITAMAGMEKGLTVIPRVVYRAFDQ-SGKR-KDNPIADIMYKGDVKKLQYRLNDRO 317
OY	378 NISGTAVYRNSRGALQLDYVPLGKGISGYFOIFQGGQSLIDYNHEATSEFGVGLMNL 437
DB	318 NYVSIVRYNPKTGYGAIEAAYTPPIKGLKGVVGRFGYGESLIDYNNHKGNGIGIGIMFN 377
OY	438 DNMGL 442

Query Match 10.2%; Score 240.5; DB 2; Length 355;
Best Local Similarity 25.1%; Pred. No. 8, 9e-12;
Matches 83; Conservative 44; Mismatches 109; Indels 95; Gaps 13;

QY 157 KDKAKOLEYAKQFTPLSLFSDLDNRNTPLMSSRPHNPMVYLPIFMGKPKRSPNTPSHE 216
DB KKYLNMDYLGTYLPLPYHSF-----TPIFQWHPHNP----- 102

QY 217 AKQTPNEFRAPELKFOVSXVKAADIMCTDSLMFQYTOOSHMOIFNGKNSRPFVYHD 276
DB 103 ---YORNEF-----KFOISFVPRVFRHILMTKGLYLYTOTNPFQIYNDPOSAFMRMIN 154

QY 277 YQPEIFLTOPYSDLPMDGKV---RMIGMAVHSNG-ESAKLSRSMNR----- 321
DB 155 FMPELIYVYPL-NKPPFGKIGNFSEIWMQIHLSNGVGAQCYQPFKEGKNPENQPFQ 213

QY 322 -----AYLMAG-----MEMK-NLTVPRIRMGRIFFEGSGSOPDD- 354
DB 214 PVIYKDYNGKQDVWGGCRSVSAGNALCFVLWMEKGGKIMVAVMPYV-----PYDQ 265

QY 355 -NPDIIDYGYGVDFLY-----OLENKSNIQGYRYRPNRSGKALQLDYVPLGKI 406
DB 266 SNPQIDIMYMGNAKIDYRGRHMFELQYDIFQWRYD--RWHGAFRLGTYRINPFV 323

QY 407 SGYFOIFQYGOQLIDYNHEATSFQVGLMN 437
DB 324 GIYAQMFNGYGDGLYEDVFSNRIGVGIRLN 354

RESULT 11

TVTRRR
protein kinase (EC 2.7.1.37) raf - raf
N:Alternate names: kinase-related transforming protein raf; raf proto-oncogene protein-s
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 11-Jun-1999
C:Accession: B26126
R:Jishikawa, F.; Takaku, F.; Nagao, M.; Sugimura, T.
Mol. Cell. Biol. 7, 1226-1232, 1987
A:Title: Rat c-raf oncogene activation by a rearrangement that produces a fused protein.
A:Reference number: A26126; MUID:87172791
A:Accession: B26126
A:Molecule type: mRNA
A:Residues: 1-602 <ISH>
A:Cross-references: GB:M15428; NID:g206546; PIDN:AAA42002.1; PID:g206547
C:Genetics:
A:Gene: raf
C:Superfamily: rat protein kinase raf; protein kinase homology
C:Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; serine/threonin
F:301-567/Domain: protein kinase homology <KIN>
F:309-317/Region: protein kinase ATP-binding motif
F:329/Active site: lys #status predicted
F:453/Binding site: phosphate (Ser) (covalent) (by autophosphorylation) #status predicted

Query Match 4.7%; Score 111.5; DB 1; Length 602;
Best Local Similarity 19.3%; Pred. No. 0.53;
Matches 93; Conservative 71; Mismatches 174; Indels 145; Gaps 23;

QY 23 QAAQVNPVAFVDEVRSENDIGONELPIDVQASASTTANPLDEHPELTTALEN 82
DB 81 ELELNNOVAVLTPKKELETAODNRNGI-----QSOFTAKEELEKEDLRT---N 131

QY 83 KTMILNCSALNODIMRLACYDTLVHGETPAVYIKTKRSIRIDETIWO--TIKGPQVYOE 140
DB 132 ERLSQEVEYLEDYKRL-NEKTESNT---TKGELQKIDELQASDVTVYKREKLEOE 186

QY 141 -----TTDPFILMGNEK-----MLTKDAQOLEYAAQFTPLS 174
DB 187 KELLHONSNMLNTELKTKTDELLALGREGKEILLKCTLENKEEDAIRSHSSASASSA 246

QY 175 LSFQDLNRNTPLMSSRPHNPMVYVLPIMHGKPKRSPNTPSHEAKQFTPNERRAPELKFOV 234
DB 247 LS-SSPNNLSTGTGWSQKTP-----VPAQREKAPSGTQEKKKIRRGGRDSYWEI 298

QY 225 SVKYAAADLNGTSDLMFQYTOOSH-----QIFNCKNSRPFVYHDYQPEI----- 281
DB 299 EASEVMSTRIGSGS---FGTVYKGMGQVAVKILKLYVDPPEQQAFFRNEVAVLRTR 355

QY 282 -----FLTO-----PYSDL-PWDGKVRI-----GMGAVAH 307
DB 356 HVNILLFMGYMTKQNLAVLTQWCESSSLXKHLHVOETFFQMFQDLIDIAQTAQGNIDYLA 415

QY 308 SNGESAKLSRSMNRAIYLAGMEK---NLTVPRIMGRIFEGSGSQ----- 351
DB 416 KNIHRRKKS--NNIFLEHGLTVKIGDEGLATVSRW-----SGSQVQEQPTGSVLM 466

QY 352 -----PDDNPDL--DYGYGVDFLYOLENKSNIQGYRYRPNRSGKALQLDYVPL 402
DB 467 APEVIRMODNPFQSDVSYSGIV-LYEL-----MTGELPYSHINRQI---IPW 515

QY 403 GKG 405
DB 516 GRG 518

RESULT 12

A29003
cellulase (EC 3.2.1.4), alkaline - Bacillus sp.
N:Alternate names: endo-1,4-beta-glucanase
C:Species: Bacillus sp.
C:Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 21-Jul-2000
C:Accession: A29003
R:Fukumori, F.; Kudo, T.; Narahashi, Y.; Horikoshi, K.
J. Gen. Microbiol. 132, 2329-2335, 1986
A:Title: Molecular cloning and nucleotide sequence of the alkaline cellulase gene fro
A:Reference number: A29003; MUID:87085443
A:Accession: A29003
A:Molecule type: DNA
A:Residues: 1-800 <PUK>
A:Cross-references: GB:D00066; GB:N00066; NID:g216223; PIDN:BA00045.1; PID:g216224
A:Experimental source: strain 1139
C:Function:
A:Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as
A:Pathway: cellulose degradation
C:Superfamily: Bacillus sp. alkaline cellulase; Thermotoga xylanase A amino-terminal
C:Keywords: glycosidase; hydrolase; polysaccharide degradation
F:565-726/Domain: Thermotoga xylanase A amino-terminal repeat homology <TXA>

Query Match 4.6%; Score 108.5; DB 2; Length 800;
Best Local Similarity 19.9%; Pred. No. 1.4;
Matches 96; Conservative 54; Mismatches 159; Indels 173; Gaps 25;

QY 8 LTLILPCFALIAIQAAQVNPVAFVDEVRSEND--LGONELPIDVQASASTDT 64
DB 10 LISSILVLVLLST-----PPTALAEGRNTRDNKRLHNDVYKRPSEAGALQLOEYDG 64

QY 65 ANPL-DEHEPELY-----TTALENKTMLNCSAL-----NODIMRLACY-DTLVHGET 110
DB 65 QMTLVDOGEGKTIQLRGMSHTGLQWPELINDNAYKALANDWESNMIRLAMYVGENGASVN 124

QY 111 PAVYIKTKRSIRIDETI-----WQT-----IKGKPOV 137
DB 125 PELIKSRVYIKGIDLAIDENDMVIYDWMYHAGCDRDPYVYAAEDFFRIALALYRNPPIIT 184

QY 138 YQETTD-----IFLMGNEKMLTKKDAQOLEYAAQFTPLSLFQDLRN---NTP 186
DB 185 YELANEPSSNNNGAGIYNNNEGMNAVYK-----EVADPIVEMLRDSGNADNIIIVGSPN 239

QY 187 WSSRP-----HNMVYLPPIF--MHGKPKNS--PNTSHEAKQFTPNERRAPELK 232
DB 240 WSQRDLAANDPIDDHHTMYTVHFYTGSHAASTESYRPETNSRGNVMSWTRVA---L 295

QY 233 QVSQYKVAEDLMGCT-----DSQDLMPFGYTOO-----SHMOIFNGKNS-----R 270
DB 296 ENGVAVPAITE--WGTSGANGCGGPFYEDADVWIEFLNENISMANNSLTN-KNEVSGAFT 352

```

QY 271 PFRVH-----DYQEIFLTQPVSDLPW---293
|:|
Db 353 PFLGKSNATSLDPPDQVWVPEELSLSGEYVARIKGVNEP---IDRKYKRVLMDFN 409

QY 294 DGKVRIGAGVAHNSGES-----AKLSRSMRATVLAGMEKKNTLVMPRIW 340
|:|
Db 410 DGKQGFV-----NGDSPVEDVYIENEGALKLSGLDASNDVSGNWMANRLSADGW 463

QY 341 GR 342
|:|
Db 464 GK 465

```

RESULT 13

cellulase (EC 3.2.1.4), alkaline - Bacillus sp. (strain KSM-64)
 N:Alternate names: endo-1,4-beta-glucanase
 C:Species: Bacillus sp.
 C>Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 10-Dec-1999
 C:Accession: J70611
 R:Sumitomo, N.; Ozaki, K.; Kawai, S.; Ito, S.
 Biosci. Biotechnol. Biochem. 56: 872-877, 1992
 A:Title: Nucleotide sequence of the gene for an alkaline endoglucanase from an alkalophilic
 A:Reference number: J70611; MUID:92305459
 A:Accession: J70611
 A:Molecule type: DNA
 A:Residues: 1-822 <SUM>
 A:Cross-references: GB:M84963; NID:g289264; PIDN:AAAT3189.1; PID:g289266
 C:Function:

A:Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as cel-
 A:Pathway: cellulose degradation
 C:Superfamily: Bacillus sp. alkaline cellulase; Thermotoga xylanase A amino-terminal reg-
 C:Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation
 F:585-726/Domain: Thermotoga xylanase A amino-terminal repeat homology <6X>

Query Match 4.6%; Score 108.5; DB 2; Length 822;
 Best Local Similarity 19.9%; Pred. No. 1.5;

Matches 96; Conservative 54; Mismatches 159; Indels 173; Gaps 25;

```

QY 8 LFLSLPCFAIIAQOAVPNVAVFVDEVRSEND---LGQDNELPIDVQASATQASSTDR 64
|:|
Db 10 LLSLIIIVLLSL-----FPTALAEQNTREDNEKHLNDNVKRPSEAGALQLOEVDG 64

QY 65 ANPL-DEHBEELY-----TTALENKTMLINCSAL-----NODIMRLACY-DLVHGET 110
|:|
Db 65 QMTLVQHGEEKIQLRGMSHTGLOMPEILINDNAVKALANDMESMTIRLAMYVGENGYASN 124

QY 111 PAVIKTKRSIRLDETT-----WQT-----IKGRQVYV 137
|:|
Db 125 PELIKSRVATKIDLALENMUYIVDMVHVAHAPGDRDPVYAGADEFFRDIALVLPNPHIT 184

QY 138 YQETTPD-----IFLMGNEKMLTKKQAKOLEYAAKQFTPLSLSFDDLRN---NTPL 186
|:|
Db 185 YELANPESSNNNGAGAPINNEEGMAVAK---ETADPIVEMLRDGNADNDITIIIGSPN 239

QY 187 WSSRP-----HNPMVLPPIF--MGKPNRS--PNTPSHEAKQFTNEFRAPELK 232
|:|
Db 240 WSGRPPLADNPIDHHTVTHVFTYGSAASTESYRPREPNSERGNVMSNTKYA---L 225

QY 233 QVSVKKAADLMGT-----DSDLMFGYTQO-----SHWQIFNGKNS---R 270
|:|
Db 296 ENGVAVFATE--WGTQANQDGGPYFDEADVWIEFLNENNISWANSILTN--KNEVSGAF 352

QY 271 PFRVH-----DYQEIFLTQPVYSDLPW---293
|:|
Db 353 PFLGKSNATSLDPPDQVWVPEELSLSGEYVARIKGVNEP---IDRKYKRVLMDFN 409

QY 294 DGKVRIGAGVAHNSGES-----AKLSRSMRATVLAGMEKKNTLVMPRIW 340
|:|
Db 410 DGKQGFV-----NGDSPVEDVYIENEGALKLSGLDASNDVSGNWMANRLSADGW 463

```

```

QY 341 GR 342
|:|
Db 464 GK 465

```

RESULT 14

cellulase (EC 3.2.1.4), alkaline - Bacillus sp. (strain KSM-S237)
 N:Alternate names: endo-1,4-beta-glucanase
 C:Species: Bacillus sp. (strain KSM-S237)
 C>Date: 31-Mar-2001 #sequence_revision 31-Mar-2001 #text_change 31-Mar-2001
 C:Accession: J7532; PC7107
 R:Hakamada, Y.; Hatake, Y.; Koike, K.; Yoshimatsu, T.; Kawai, S.; Kobayashi, T.; Ito,
 Biosci. Biotechnol. Biochem. 64, 2281-2289, 2000
 A:Title: Deduced amino acid sequence and possible catalytic residues of a thermostabl
 A:Reference number: J7532; MUID:21036886
 A:Accession: J7532
 A:Molecule type: DNA
 A:Residues: 1-824 <HAK>
 A:Cross-references: DDBJ:AB018420
 A:Experimental source: strain KSM-S237
 A:Accession: PC7107

A:Molecule type: protein
 A:Residues: 31-50 <HAZ>
 C:Comment: This enzyme is a thermostable, alkaline enzyme used as effective detergent
 C:Genetics:
 A:Gene: Egl-237
 C:Keywords: hydrolase; glycosidase

Query Match 4.5%; Score 107.5; DB 3; Length 824;
 Best Local Similarity 19.7%; Pred. No. 1.8;

Matches 91; Conservative 56; Mismatches 147; Indels 167; Gaps 27;

```

QY 30 PAVFVDEVRSEND---LGQDNELPIDVQASATQASSTDRANPL-DEHBEELY-----TT 78
|:|
Db 25 PALLAEQNTREDNEKHLNDNVKRPSEAGALQLOEVDGQMTLVQHGEEKIQLRGMSHT 84

QY 79 ALENKTMILINCSAL-----NODIMRLACYDLVHGE-----TPAVIKTK-----RSI 120
|:|
Db 85 GLQMFPEILINDNAVKALNSDMSMTIRLAMY-----VGENGYANPELIKQRYVIDGIELAI 140

QY 121 RLDETT---WQT-----IKGRQVYVQETTPD-----IF 146
|:|
Db 141 ENDMUYIVDMVHVAHAPGDRDPVYAGAKDFPRETALVLPNPHITIELANPESSNNNGAG 200

QY 147 LMGNEKMLTKKQAKOLEYAAKQFTPLSLSFDDLRN---NTPLMSRP-----H 192
|:|
Db 201 IPNNEBGMKAUK---ETADPIVEMLRKSGNADNDITIIIGSPWSORPDLADNPIDH 255

QY 193 NPMVLPPIFHMGRPNRSPNTPSHEAKQFTNEFRAPELK-----FQVSVKKAADLMGT 247
|:|
Db 256 HTMTVTHVFTYGSAASTESYRPREPNSERGNVMSNTKYA---WGT 308

QY 248 -----DSDLMFGYTQO-----SHWQIFNGKNS---RPRV-----274
|:|
Db 309 SQASDGGPYFDEADVWIEFLNENNISWANSILTN--KNEVSGAFPELCKSNATNLDPG 367

QY 275 --HDYQF-EFLF-----QPV-----YSDLPW---DGKVMIGAGVAHNSGE 311
|:|
Db 368 PDHVAPEELSLSGEYVARIKGVNEPDIRTYTKLWLFQNGTQO--GFGVNSDSPNK 425

QY 312 SAKLSRSMRATVLAGME-----WKNLTVMPIWGR 342
|:|
Db 426 ELIAYDENNTLVKVSGLDVSNVSDGNFMANRLSANGWCK 466

```

RESULT 15

hypothetical protein F25B3.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T20122; T21324

This Page Blank (uspto)

Query Match 15.2%, Score 358, DB 1, Length 289;

OY		226	RAPELKFOVSVKVAKAEDLMGTDSLMLMGCTYOOSHQMFNGNNSRPFVRHDXQPELFLIQ	285
Db		83	KRDVEKFOLFSLIAFPLMRGILGPNSVLGASYOKSMWOLSNSESSPFRFTNEPOLFGF	142
OY		286	PV-YESDLPWDGCVRRIGCGAVAHNSGESAKLSRSNNRAYLMAGMKMKLTVPRIMGRIF	344
Db		143	ATDYFAACM--TLRVHEVGNYHNDSNGRSDPTSRITNNRLYLTRLAENGNMLVFKRMYYI-	199
OY		345	KEGSGSPDDNDPDLLIDYGYGDVREFLYOLENKSNISGTYRVNPRSGKALQLDYVPYLGK	404
Db		200	-----GSTDDMDNDIKRYMGYOLKIGHYHL-GEAVYSAGKYWMNTGYGGAEGLSPYTK	253
OY		405	GISGYFOJFOGCGOSLIDYNHEATSFEGCLMIND	438
Db		254	HVRLTYOVSYGEGSELIDNFNQIRVGVGMLEND	287
RESULT		3		
ID	PAL_ECOLI	STANDARD:	PRT:	289 AA.
AC	P00631.			
DT	21-JUL-1986	(Rel. 01, Created)		
DT	15-JUL-1998	(Rel. 36, Last sequence update)		
DT	20-AUG-2001	(Rel. 40, Last annotation update)		
DE	PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DEREGENT-RESISTANT			
DE	PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYCHOLINE 1-			
DE	ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).			
GN	PLDA OR PB821 OR Z5342 OR ECS4751.			
OS	Escherichia coli, and			
OS	Escherichia coli O157:H7.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
CC	Escherichia.			
OX	NCHI_TaxID=562, 83334;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	MEDLINE=85157492; PubMed=6397464;			
RA	Homma H., Kobayashi T., Chiba N., Karasawa K., Mizushima H., Kudo I.,			
RT	Inoue K., Ikeda H., Sekiguchi M., Nojima S.;			
RT	"The DNA sequence encoding plidA gene, the structural gene for			
RL	detergent-resistant phospholipase A of E. coli.";			
RN	J. Biochem. 96:1655-1664(1984).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-K12 / MG1655;			
RA	MEDLINE=92358234; PubMed=1379743;			
RT	Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;			
RT	"Analysis of the Escherichia coli genome: DNA sequence of the region			
RL	from 84.5 to 86.5 minutes.";			
RN	Science 257:771-778(1992).			
RN	[3]			
RP	REVISION TO 14-15.			
RC	STRAIN-K12 / MG1655;			
RA	MEDLINE=97426617; PubMed=9278503;			
RA	Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,			
RA	Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,			
RA	Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,			
RA	Mau B., Shao Y.;			
RT	"The complete genome sequence of Escherichia coli K-12.";			
RL	science 277:1453-1474(1997).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-O157:H7 / EDL933 / ATCC 700927;			
RA	MEDLINE=21074935; PubMed=11206551;			
RA	Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,			
RA	Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,			
RA	Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,			
RA	Grotbeck E.J., Davis N.W., Lim A., Dimantanta E.T., Potamousis K.,			
RA	Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,			
RA	Weich R.A., Blattner F.R.;			
RT	"Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";			

RL Nature 409:529-533(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / RIMD 0509952;
RX MEDLINE-21156231; PubMed-11258796;
RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Mureta T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
RA Kihara S., Shiba T., Hattori M., Shiga H.,
RT "Complete genome sequence of enterohemorrhagic *Escherichia coli*
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [6]
RP SEQUENCE OF 30-289 FROM N.A.
RC STRAIN-K12;
RX MEDLINE-85003590; PubMed-6383820;
RA de Gans P., Verheij H.M., Riegman N.H., Hoekstra W.P.M., de Haas G.H.,
RT "The pro- and mature forms of the *E. coli* K-12 outer membrane
RT phospholipase A are identical.";
RL EMBO J. 3:1799-1802(1984).
RN [7]
RP SEQUENCE OF 174-289 FROM N.A.
RC STRAIN-K12;
RX MEDLINE-87115164; PubMed-3027506;
RA Iritano N., Nakayama K., Nakayama H.,
RT "The recQ gene of *Escherichia coli* K12: primary structure and
RT evidence for SOS regulation.";
RL Mol. Gen. Genet. 205:298-304(1986).
RN [8]
RP MUTAGENESIS OF SER-172.
RX MEDLINE-94131966; PubMed-8300539;
RA Brok R.G.P.M., Brinkman E., van Hoxtel R., Bekkers A.C.A.P.,
RA Verheij H.M., Tommassen J.,
RT "Molecular characterization of enterobacterial pldA genes encoding
RT outer membrane phospholipase A.";
RL J. Bacteriol. 176:861-870(1994).
RN [9]
RP ACTIVE SITE SER-164, AND PARTIAL SEQUENCE.
RX MEDLINE-91249806; PubMed-2040286;
RA Horevoets A.J.G., Verheij H.M., de Haas G.H.,
RT "Inactivation of *Escherichia coli* outer-membrane phospholipase A by
RT the affinity label hexadecanethionyl fluoride. Evidence for an
RT active-site serine.";
RL Eur. J. Biochem. 198:247-253(1991).
CC -1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE
CC A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.
CC REQUIRED FOR EFFICIENT SECRETION OF BACTERIOCINS, SEEMS TO BE
CC DOMINANT IN NORMAL GROWING CELLS.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.
CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES
CC LOCATED THERE.
CC -1- INDUCTION: BY MEMBRANE DAMAGING, FOR EXAMPLE, BY PHAGE-INDUCED
CC LYSIS OR TEMPERATURE SHOCK.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on les-
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL, X02143; CAA26081.1; -;
DR EMBL, M87049; AAA67617.1; -;
DR EMBL, AE000458; AAC76824.1; -;
DR EMBL, AE005613; AAG59017.1; -;
DR EMBL, AP002567; BAB38174.1; -;
DR EMBL, M30198; AAA24516.1; -;
DR PIR, A00771; PSECA.

[illegible]


```

RESULT 8
GUN_PAPEO STANDARD: PRT: 397 AA.
AC P23548:
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE ENDOGUCANASE (EC 3.2.1.4) (ENDO-1,4-BETA-GLUCANASE) (CELLULOSE).
OS Paenibacillus polymyxa (Bacillus polymyxa).
OC Bacteria: Firmicutes: Bacillus/Clostridium group:
CC Bacillus/Staphylococcus group: Paenibacillus.
OX NCBI_Taxid=1406;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90170877; PubMed=2307659;
RA Baird S.D., Johnson D.A., Selligy V.L.;
RT "Molecular cloning, expression, and characterization of
  endo-beta-1,4-glucanase genes from Bacillus polymyxa and Bacillus
  circulans."
RL J. Bacteriol. 172:1576-1586(1990).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
  LINKAGES IN CELLULOSE.
CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY A (FAMILY 5 OF GLYCOSYL
  HYDROLASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M33791; AAA22631.1; -
DR PIR: A35136; A35136.
DR HSSP: P54383; 1ECE.
DR InterPro: IPR001547; Glyco_hydro_F5.
DR Pfam: PF00150; cellulase.1.
DR PROSITE: PS00659; GLYCOSYL_HYDROL_F5; 1.
KW Cellulose degradation: Hydrolase: Glycosidase.
FT ACT_SITE 194 194 PROTON DONOR (BY SIMILARITY).
FT ACT_SITE 317 317 NUCLEOPHILE (BY SIMILARITY).
SQ SEQUENCE 397 AA; 44357 MW; B9C2E802C04F0A2A CRC64;

Query Match 4.0%; Score 94.5; DB 1; Length 397;
Best Local Similarity 25.7%; Pred. No. 3;
Matches 46; Conservative 21; Mismatches 61; Indels 51; Gaps 10;

QY 301 GNGAVHNGESAKLSRSNNRAYLMAGMEKMLTVMPRTMGR-----IFKEG----- 347
DB 43 GKRIYDESKKEAFNGLNM-----FGLETPNPT-LHGMSRMDMDLQVKEGYNLR 95
QY 348 -----SGSQPDD-----NPDILDYGYGVDFRFLYLENKSNTSGTV---RYNPRS 389
DB 96 LPYSNOLFSSSRPDSIDYKKNPDV---GLNPIQIMDKLEKAGRGIOIILDRHRPS 152
QY 390 GNGAQLDLV--YPLGKGISGYFQIFQGY-----GSLIDYNHNAITSEGVGLMLNDW 439
DB 153 G-GQSELWYTSQYPSRWISDMKMLADRYKKNMFTYIGADLHNPHQASMGTCGNASTDW 210

RESULT 9
BCK1_YEAST
ID BCK1_YEAST STANDARD: PRT: 1478 AA.
AC Q01389; P32894;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SERINE/THREONINE PROTEIN KINASE BCK1/SLK1/SSP31 (EC 2.7.-.-).
GN BCK1 OR SLK1 OR SSP31 OR LMS3 OR YJL095W OR J0906.

```

```

OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_Taxid=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92186847; PubMed=1545797;
RA Costigan C., Gehring S., Snyder M.;
RT "A synthetic lethal screen identifies SLK1, a novel protein kinase
  homolog implicated in yeast cell morphogenesis and cell growth."
RL Mol. Cell. Biol. 12:1162-1178(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92104496; PubMed=1840547;
RA Irie K., Araki H., Oshima Y.;
RT "A new protein kinase, SSP31, modulating the SMP3 gene-product
  involved in plasmid maintenance in Saccharomyces cerevisiae."
RL Gene 108:139-144(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=EG123;
RA MEDLINE=92107166; PubMed=1729597;
RA Lee K.S., Levin D.E.;
RT "Dominant mutations in a gene encoding a putative protein kinase
  (BCK1) bypass the requirement for a Saccharomyces cerevisiae protein
  kinase C homolog."
RL Mol. Cell. Biol. 12:172-182(1992).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C;
RA MEDLINE=95176706; PubMed=7871887;
RA Miosga T., Boles E., Schaaff-Gerstenschlaeger I., Schmitt S.,
  Zimmermann F.K.;
RT "Sequence and function analysis of a 9.74 kb fragment of
  Saccharomyces cerevisiae chromosome X including the BCK1 gene."
RL Yeast 10:1481-1488(1994).
RN [5]
RP SEQUENCE OF 602-1104 FROM N.A.
RA Cusick M.E.;
RT Submitted (XXX-1992) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: SERINE/THREONINE PROTEIN KINASE INVOLVED IN A SIGNAL
  TRANSDUCTION PATHWAY THAT PLAY A ROLE IN YEAST CELL MORPHOGENESIS
  AND CELL GROWTH. THIS PATHWAY SEEMS TO START BY SMP3, THEN
  INVOLVE THE KINASE PC1 THAT MAY ACT ON THIS KINASE. BCK1 PROBABLY
  PHOSPHORYLATES MKR1 AND MKR2 WHICH THEMSELVES PHOSPHORYLATE THE
  MPK1 KINASE.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M84389; -; NOT_ANNOTATED_CDS.
DR EMBL: D10389; BAA01226.1; -
DR EMBL: X60227; CAA42788.1; -
DR EMBL: X77923; CAA54896.1; -
DR EMBL: Z49370; CAA89389.1; -
DR EMBL: Z49369; CAA89388.1; -
DR EMBL: M88604; AAA21179.1; -
DR PIR: S20117; S20117.
DR PIR: S22285; S22285.
DR PIR: J01118; J01118.
DR PIR: J01432; J01432.
DR HSSP: P24941; 1A01.
DR SGD: S0003631; BCK1.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_kin_actsite.

```

DR Pfam: PF00069; kinase: 1.
 DR SMART: SM00220; S_TKC: 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP: 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST: 1.
 DR TRANSFERASE: P550011; PROTEIN_KINASE_DOM: 1.
 KW ATP-binding; Phosphorylation.
 KW ATP-binding; Phosphorylation.
 FT DOMAIN 1175 1440
 FT NP_BIND 1181 1189
 FT BINDING 1204 1204
 FT ACT_SITE 1303 1303
 FT MOD_RES 1134 1134
 FT VARIANT 1119 1119
 FT VARIANT 1120 1120
 FT VARIANT 1120 1120
 FT VARIANT 1146 1146
 FT VARIANT 1174 1174
 FT VARIANT 59 59
 FT CONFLICT 79 79
 FT CONFLICT 264 264
 FT CONFLICT 279 279
 FT CONFLICT 703 714
 FT CONFLICT 795 795
 FT CONFLICT 802 802
 FT CONFLICT 808 808
 FT CONFLICT 903 903
 FT CONFLICT 919 919
 FT CONFLICT 960 962
 FT CONFLICT 1086 1104
 SQ SEQUENCE 1478 AA; 164194 MW; D586C3A497A5BB33 CRC64;
 H (IN REF. 5).
 RPPVDSSEYFEIDGLNGK -> VPIAHTSYRMDLVKIN
 (IN REF. 5).
 ADA -> RDR (IN REF. 3).
 T -> N (IN REF. 3).
 T -> N (IN REF. 3).
 A -> S (IN REF. 3).
 L -> V (IN REF. 3).
 S -> A (IN REF. 3).
 RYQPSYYDR -> STPKRVITMT (IN REF. 3).
 N -> I (IN REF. 3).
 A -> P (IN REF. 3).
 E -> V (IN REF. 3).
 A -> I (IN REF. 3).
 F -> I (IN REF. 3).
 A -> P (IN BCK1-20; ACTIVATION).
 G -> V (IN BCK1-10; ACTIVATION).
 I -> T (IN BCK1-16; ACTIVATION).
 T -> K (IN BCK1-11; ACTIVATION).
 T -> P (IN BCK1-19; ACTIVATION).
 PHOSPHORYLATION (BY PKC) (POTENTIAL).
 BY SIMILARITY.
 ATP (BY SIMILARITY).
 ATP (BY SIMILARITY).

Query Match 4.0%; Score 94.5; DB 1; Length 1478;
 Best Local Similarity 22.5%; Pred. No. 19;
 Matches 98; Conservative 51; Mismatches 156; Indels 131; Gaps 24;

OY 36 EVKSENDLGQNDLPI--DVQASATQASTDTANPLDEHEPELYTALENK--TMLINCS 90
 DB 950 DIFEENDITFADAPPFDDSDDDSSDDIWSKKKTAPE--TNNENKDEKSDNS 1006
 OY 91 ALNODIMRLACYDTLVHGEFTPAVITKRSIRLDETIMOTIKGPOVVOETDPIRLMN 150
 DB 1007 THSDEIF--YDS--QTQKMERK-----MTFRSPPEVVO-NLEKFPFRAN 1047
 OY 151 EKGMLTRKDAQLEVAQKQFPLSLFDDLRNNTPLMSSRPNNPVYDLPIFMHGFNRS- 209
 DB 1048 LDKPIT-----EGIASPTSPKSLDSLSPKNVASSRTEPSPRPV-----PDSSY 1094
 OY 210 -----PNTPSHEAKQFTNERRA-----PELAFQVSVKKAAL-DLMGTDSDIM 252
 DB 1095 EFIDGLNGKKNPLNOAK--TEKRTKRTITTAHEASLARKNSVKLKRONTKMWGT---RM 1149
 OY 253 FGYTQOASHMOJENGNKNSRFVHDYQPEIFLQPYSDLPMDGKVMIMGA-----V 305
 DB 1150 VEVTEHNHWSIKAKNSK-----GEYKEPAM-MKGEMIKGSGGAYVLCI 1193
 OY 306 HHSNESEAKSLR-----SMNRAYLMAGMEKKNLTVMPRIRGIRFKEGSGSQPDNDPI 358
 DB 1194 NVTTEGMAVVKQVEPKYSSONEALLS-----TV-----EALHSEVSTLKLDHLMNI 1240
 OY 359 LDYVGYGVRFILYOLENSNIGSYVRYNPRSGKALQLDYVPLKGIGSYQIPIFGYGO 418
 DB 1241 VOYIGF-----ENKNNTY-----SLFELEYV--AGSGVSLIRMGRFDE 1277
 OY 419 SLIDYNHEATSPGVGL 434
 DB 1278 PLI--KHLTYOVLKGL 1291
 RESULT 10
 SRK6_BRAOL STANDARD: PRT: 849 AA.

AC 009092;
 DT 01-FEB-1995 (rel. 31, Created)
 DT 01-FEB-1995 (rel. 31, Last sequence update)
 DT 01-NOV-1995 (rel. 32, Last annotation update)
 DE PUTATIVE SERINE/THREONINE KINASE RECEPTOR PRECURSOR (EC 2.7.1.37)
 DE (S-RECEPTOR KINASE) (SRK).
 GN SRK6.
 OS Brassica oleracea (Cauliflower).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Brassica.
 ON NCBI_Taxid=3712;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. S656; TISSUE=Stigma;
 RC MEDLINE=92020942; PubMed=1681543;
 RX Stein J.C., Howlett B., Boyes D.C., Nasrallah M.E.;
 RA "Molecular cloning of a putative receptor protein kinase gene encoded
 RT at the self-incompatibility locus of Brassica oleracea.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:8816-8820(1991).
 CC -1- FUNCTION: INVOLVED IN SPOROPHYTIC SELF-INCOMPATIBILITY SYSTEM
 CC (THE INABILITY OF FLOWERING PLANTS TO ACHIEVE SELF-
 CC FERTILIZATION), PROBABLY ACTING IN COMBINATION WITH S-LOCUS-
 CC SPECIFIC GLYCOPROTEINS. INTERACTION WITH A LIGAND IN THE
 CC EXTRACELLULAR DOMAIN TRIGGERS THE PROTEIN KINASE ACTIVITY OF THE
 CC CYTOPLASMIC DOMAIN.
 CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN -> ADP + A PHOSPHOPROTEIN.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY IN THE PISTIL AND ANTHER.
 CC -1- POLYMORPHISM: THERE ARE A NUMBER OF DIFFERENT S ALLELES IN
 CC B.OLERACEA, POSSIBLY PROVIDING THE RECOGNITION SPECIFICITY.
 CC -1- SIMILARITY: THE EXTRACELLULAR DOMAIN IS SIMILAR TO S-LOCUS
 CC GLYCOPROTEINS OF BRASSICA, WHILE THE INTRACELLULAR DOMAIN IS
 CC A SER/THR-PROTEIN KINASE RELATED TO RAF KINASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: M76647; AAA33000.1; ALT-TERM.
 DR HSSP: P11362; IFGI.
 DR InterPro: IPR001480; B_lectin.
 DR InterPro: IPR000719; Euk-kinase.
 DR InterPro: IPR003609; Pan-app.
 DR InterPro: IPR002280; Ser_thr_kin_actsite.
 DR InterPro: IPR000858; Slocus_glycop.
 DR Pfam: PF00069; kinase: 1.
 DR Pfam: PF00954; S_locus_glycop: 1.
 DR SMART: SM00108; B_lectin: 1.
 DR SMART: SM00473; PAN_AP: 1.
 DR SMART: SM00221; STYKC: 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP: 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST: 1.
 DR TRANSFERASE: P550011; PROTEIN_KINASE_DOM: 1.
 KW Transmembrane; Receptor; Glycoprotein; Self-incompatibility.
 FT SIGNAL 1 32
 FT CHAIN 33 849
 FT FT
 FT DOMAIN 33 446
 FT TRANSMEM 447 446
 FT DOMAIN 467 466
 FT DOMAIN 528 779
 FT NP_BIND 534 542
 FT BINDING 556 556
 FT ACT_SITE 653 653
 FT CARBOHYD 47 47
 FT CARBOHYD 120 120
 FT CARBOHYD 196 196
 N-TERMINAL
 PUTATIVE SERINE/THREONINE KINASE
 RECEPTOR.
 EXTRACELLULAR (POTENTIAL).
 POTENTIAL.
 CYTOPLASMIC (POTENTIAL).
 PROTEIN KINASE
 ATP (BY SIMILARITY).
 ATP (BY SIMILARITY).
 BY SIMILARITY.
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 389 389 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 849 AA: 97231 MW: 7E156059EEDDF4370 CRC64;

Query Match 4.0%; Score 93.5; DB 1; Length 849;
 Best Local Similarity 21.9%; Pred. No. 10;
 Matches 78; Conservative 50; Mismatches 107; Indels 121; Gaps 24;

QY 105 LVHGEPAVVKTKRSINLDEIWMOTIKGPOVVOETDPIELMG---NKK---GMLTK 157
 DB 24 LHPALSIYNT---LSSTESLTISSNKTLV---SPGSEFEVGFPRTRNSRWLGMYWK 75
 QY 158 K-DAKOLEYAKOFTPLSLSD-----LDRNNTPLMSSRPNNPVMVLEPIFMHGK 205
 DB 76 KVSDFRTVWVANRNPPLSNMIGTLKISGNNLVLDHSHKPKVWMTN-----LIRGN 125
 QY 206 PNRSP-----NTPSHEAKOFT-----PNEFRAPELKFOVSVKRAEDLW 245
 DB 126 -ERSPVVAELIANGNFVWRDSSNNDASEYLMQSEFDYPTDILLPEMKLGYNLKT----- 177
 QY 246 GTDSDLMEFGYTQOSHMQIFNCKNSRPRVHDYO-----PEIFLQO---PYVSDLPMDGK 296
 DB 178 GLNFFL-----TSMSSDDPSSGNF---SYKLETSQSLPEFYLSRENEFPHRSGPMWG- 226
 QY 297 VRMIGMAVH-----HSNGESAKLSRNMRAV-----LMAGMEKMLTYMP--RI 339
 DB 227 IREFGIEDQKLSYMYNFIENNEVAYTFRTMNSPYSRLTSLSEGFQGLTWYPSIRI 286
 QY 340 WGRIFKSGSGQPDNDPILDY-----YGYDVRFLYOLENKSNSIGTVR-YNPRS 389
 DB 287 WNRFW-----SSPYD-POCDTYIMCGPYAYCDV-----NTSPVCNCIOGFNPRN 329

RESULT 11
 VG14_BPML5 STANDARD; PRT; 486 AA.
 AC 0052220;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE GENE 14 PROTEIN (GP14).
 GN 14.
 OS Mycobacteriophage L5.
 OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae.
 OX NCBI_TaxID=31757;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93211282; Pubmed=8459766;
 RA Hatfull G.F., Sarkis G.J.;
 RT "DNA sequence, structure and gene expression of mycobacteriophage L5:
 RT a phage system for mycobacterial genetics.";
 RL Mol. Microbiol. 7:395-405(1993).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: Z18946; CAA79390.1; -
 DR PIR: S30959; S30959;
 DR SEQUENCE 486 AA: 53757 MW: 4BF37346BB11BF5 CRC64;

Query Match 3.9%; Score 92.5; DB 1; Length 486;
 Best Local Similarity 21.0%; Pred. No. 5.8;
 Matches 99; Conservative 69; Mismatches 173; Indels 131; Gaps 24;

QY 21 IQQAQA-VPNVAVDEVRSN-----DLGDNELPIDVQATQSASTDTANPDEHEPE 74
 DB 56 MOQLAHGVPRLVYVAERQAVEGRFLGDADADELMQWQ-----ANNIDIAEP 109
 QY 75 LYTALEKMTKLINCSA-----LNODIMRLACVDTLVHGE-TPAVIKTRSIIRD 123
 DB 110 GYDAVYHGRSEFTITISKPDPLDGLMDONVPIIIVE-PTTMHAELIDPRINRSKAIR-- 166
 QY 124 ETIMQTIKGRPOVVOETDPIELMGNEKMLTKDAKOLEYAKOFTPL-SLSFDLDRN 182
 DB 167 -----VAYDKR-----GNE-----IQAAATLYPMETIGFRADG 195
 QY 183 NTPLMSSRPNNPVMVLEPIFMHGKPNRSPNTPSHEAKOFTPEFRAPELKFOVSVKRAAE 242
 DB 196 EWAEWFNVPHG-LGVVPVY--PLPNRRLSDLYGTSETPELRSMQTAARILMLQATA 252
 QY 243 DLMEGDSDLME-----GYTQOSHMQIFNCKNSRPRVHDYOPEI-----F 282
 DB 253 ELKGVPRQLIGIRPEELGVSETGQTLFDAYLRIALFEDEAGKIQOFSAAELANFTNA 312
 QY 283 LTQ-----PYVSDLPMDGKVRMIGMAVHSHNGES-----AKLSRNMRAVLMAGMEKN 332
 DB 313 LDQIAKQVAAVTGLP---POLYSTAADNPASAEAIRAESRLIKYERKNLMFGAMEE 368
 QY 333 LTVMPRIKGRIFKSGSGQPDNDPILDYGYGCVRFYOLENKSNSIGTVRNPBSGK 392
 DB 369 AM---RIAYRIMKG---DVPDMT-----RME-----TYWRDPSTPTTY 401
 QY 393 ALQD-----YVYPLGKI---SGYFOIFQGYG-----QSLIDYHNEATSFVGILM 435
 DB 402 AAKADAAKTLKINGOGVTPRRARIDMGYSYKEREEMRMBEAAKMLGILL 453

RESULT 12
 BAC2_MOUSE STANDARD; PRT; 716 AA.
 AC P97303;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE TRANSCRIPTION REGULATOR PROTEIN BACH2 (BTB AND CNC HOMOLOG 2).
 GN BACH2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BA1B/C;
 RX MEDLINE=97042438; Pubmed=8887638;
 RA Oyake T., Itoh K., Motohashi H., Hayashi N., Hoshino H., Nishizawa M.,
 RA Yamamoto M., Igarashi K.;
 RT "Bach proteins belong to a novel family of BTB-basic leucine zipper
 RT transcription factors that interact with Mafk and regulate
 RT transcription through the NF-E2 site.";
 RL Mol. Cell. Biol. 16:6083-6093(1996).
 CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS AS REPRESSOR OR
 CC ACTIVATOR. BINDS, IN-VITRO, TO NF-E2 BINDING SITES. PLAY IMPORTANT
 CC ROLES IN COORDINATING TRANSCRIPTION ACTIVATION AND REPRESSION BY
 CC MAFK.
 CC -1- SUBUNIT: HETERODIMER OF BACH2 AND MAFK.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSION RESTRICTED TO MONOCYTES AND
 CC NEURONAL CELLS.
 CC -1- SIMILARITY: BELONGS TO THE BTB/POZ DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 BTB/POZ DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

RC TISSUE-Breast: PubMed-7641693:
RX MEDLINE-95369246:
RA Cavailles V., Dauvois S., L'Horsset F., Lopez G., Hoare S.,
RA Kusner P.J., Parker M.G.:
RT "Nuclear factor R1P40 modulates transcriptional activation by the
RT estrogen receptor".
RL EMBO J. 14:3741-3751(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-20289799; PubMed-10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishi K., Totoki Y., Choi D.-K., Soeda E.,
RA Ohki M., Takagi T., Sakaki Y., Taudien S., Bleeschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyma S., Antonarakis S.E.,
RA Mitsuima S., Shintzu N., Norisiek G., Hornischer K., Brandt P.,
RA Scharte M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E.,
RA Weinmeyer S., Borzym K., Gardiner K., Mizetic D., Francis F.,
RA Lehnach H., Reinhardt R., Yaspo M.-L.:
RT "The DNA sequence of human chromosome 21."
RL Nature 405:311-319(2000).
CC -1- FUNCTION: MODULATES TRANSCRIPTIONAL ACTIVATION BY THE ESTROGEN
CC RECEPTOR.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X84373; CAA5108.1; -
DR EMBL: AL163207; CAB90396.1; -
DR MIM: 602490; -
KM Transcription regulation; Nuclear protein.
FT CONFLICT 124 124 P -> R (IN REF. 1).
FT CONFLICT 448 448 R -> G (IN REF. 1).
FT CONFLICT 721 726 NKGSE -> TKGRVK (IN REF. 1).
FT CONFLICT 1080 1080 T -> A (IN REF. 1).
SQ SEQUENCE 1158 AA; 126941 MW; 81RC424968B9A5F6 CRC64;

Query Match 3.9%; Score 92; DB 1; Length 1158;
Best Local Similarity 20.0%; Pred. No. 21;
Matches 93; Conservative 62; Mismatches 138; Indels 172; Gaps 23;

OY 41 NDLOGDNELPIDVGSATOSASIDTANPLDEHEPELYTTA-----LEKTYMLIN----- 88
DB 524 NDVSFNFQNYARISVSISSPTNRTTPVS--TPPLTSSKAGSPINLSQSHSLVIRWNSPP 581
OY 89 --CSALNDIMRLACYDTL-----VHGEPYAVIK-----TKRSIRDETIMOTIKG 133
DB 582 YVCSQSEKLTNTASNSHMDLTKSKDPPEKPAONEGAMNSTFASKILQWLAOCQMOS 641
OY 134 PQVYQOETTDPIFLMGNEKGMILTKDAKOLEYAAKQFTPLSLSPDLDRNNRPL----- 186
DB 642 SMSVEGRPSKQLLTGN-----TDK-----PIGM-----IDRLNSPLSNKNTNA 680
OY 187 -----WSSRPINP-----MYVLPFPMHGKPRKSRNNTSHEKQTP--- 222
DB 681 VEENKAFFSQPGPPEGLSGSEIENLLERTYLOLLT-GNPKGKS-----EKKEETPLRD 735
OY 223 -----NEFRAPELKFOYSKYVKAEDLWGTDSDLMFGYTGQSHQIFNGKNSRPF----- 272
DB 736 ESTGHSERALSQILMKYIKISEPCDDLOIPNTV-----HLSHDAKSAPFLGMA 785
OY 273 -RVHDYQP-----ELFLQPVYSDLPMDGKVMIGMGAHVHNSGESAKLRSWNRAYL-- 324

DB 786 PAYORSAPALPYSEDFKSEPV---SPQDFS-----FSKNGLLSLRLQNDOSYLAD 833
OY 325 -----MAGMKWULTYMPRIMGRIFKEGSGSQDDNP-----DIIDYIG----- 363
DB 834 DSDSRHNNENALLSKNLCVWP-----KRKLTEBLENPFKKMKNNIVDAANNHAP 887
OY 364 ---YGDVRFVLEENKNSINSVTRNPRSGALQLDVYVPLGKG 405
DB 888 EYLSGLNLOEBLFRN-----DLERKYPAGHG 916

RESULT 15
CATA.DICDI STANDARD: PRT: 496 AA.
ID CATA.DICDI
AC 077229;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CATALASE (EC 1.11.1.6).
GN CATA OR CAT.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RA Foote C., Alexander H., Alexander S.;
RL Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE.
CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC -1- CORFACTOR: HEME GROUP (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF090443; AAC36743.1; -
DR HSSP: P00432; 7CAT.
DR DictyDb: DD07777; cata.
DR InterPro: IPR002226; Catalase.
DR Pfam: PF00199; catalase.1.
DR PRINTS: PR00067; CATALASE.
DR ProDom: PD000510; Catalase.1.
DR PROSITE: PS00437; CATALASE_1; 1.
DR PROSITE: PS00438; CATALASE_2; 1.
KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Peroxisome.
FT ACT_SITE 54 54 BY SIMILARITY.
FT BINDING 128 128 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT SITE 338 338 MICROBODY TARGETING SIGNAL (POTENTIAL).
FT SITE 494 496
SQ SEQUENCE 496 AA; 55683 MW; 683318B8BARD2E6 CRC64;

Query Match 3.9%; Score 91; DB 1; Length 496;
Best Local Similarity 23.5%; Pred. No. 7.9;
Matches 56; Conservative 32; Mismatches 94; Indels 56; Gaps 12;

OY 61 STDTPANPLDEHEPELYTTALEKTYMLINCSALNODIMRLACYDTLVHGEPYAVIKTRSI 120
DB 7 TTSSGSPIDNN-----LNSMTAGVNGPILIDQFTLI-----DKLAHNDRE----- 46
OY 121 RIDETIMOTIKGPQVYQOETTD-----PIF-----LMGNEKGM-L 156
DB 47 RIFERVHA-KGAGAHGFVEYTSDDVPKMKAKFLNKVGKRPITFTSTVGEGKSSDS 105

Qy 157 KKDAGQLEYAKQFTPLSLSEDLDRNNTPLMSSRPINPMYVLPITFMHGKPNRSPNTPSHE 216
Db 106 ERDPR-GFAVKFYEEG-NEDWVGNNTPVFEIIRDPK--FPDFIHTQ-KRNPQTNCCKD 158
Qy 217 AKQFTPNFRAPELKFQVSVKVAEDLMGTDSDLWFGYTOOSHWOIFNGKNSRPFV 274
Db 159 PNMFMDFLGOTPESTHQVSLFSDR---GTPKSYRHHMGFSHTLKFVNAQCKPYW 212

Search completed: November 30, 2001, 14:27:06
Job time: 553 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 13:05:18 ; Search time 77.39 Seconds
(without alignments)
835.410 Million cell updates/sec

Title: US-09-787-083-8
Perfect score: 442
Sequence: 1 MKVSLSTLTLSTLPCFALLA.....YHNEATSFVGLMLDMWML 442

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 473505 seqs, 146272329 residues

Word size : 0

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPREMBL.17:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	2.0	382	2 09K07	09K07 neisseria m
2	9	2.0	409	2 09J721	09J721 neisseria m
3	9	2.0	862	2 09CL62	09CL62 pasteurella
4	9	2.0	1432	10 09AX85	09AX85 oryza sativ
5	8	1.8	299	10 09M268	09M268 arabidopsis
6	8	1.8	306	2 09CL22	09CL22 pasteurella
7	7	1.6	55	12 067608	067608 tomato gold
8	7	1.6	72	12 0912S5	0912S5 potato viru
9	7	1.6	72	12 0912S4	0912S4 potato viru
10	7	1.6	95	12 0912S3	0912S3 potato viru
11	7	1.6	96	12 0912S8	0912S8 potato viru
12	7	1.6	101	12 0912S6	0912S6 potato viru
13	7	1.6	122	2 09R016	09R016 delinococcus
14	7	1.6	134	12 0912S7	0912S7 potato viru
15	7	1.6	131	6 09GMD4	09GMD4 macaca fusc
16	7	1.6	138	12 0912Q7	0912Q7 potato viru
17	7	1.6	147	5 036035	036035 hexamita in
18	7	1.6	158	2 0929C1	0929C1 chlamydia p
19	7	1.6	169	2 0912A7	0912A7 streptomyces

20	7	1.6	186	10 09LRX9	09LRX9 arabidopsis
21	7	1.6	189	13 09DFT2	09DFT2 notochonia
22	7	1.6	198	4 09BUX9	09BUX9 homo sapien
23	7	1.6	212	13 091819	091819 ictalurus p
24	7	1.6	214	5 09VTR5	09VTR5 drosophila
25	7	1.6	216	2 09X6H6	09X6H6 streptococ
26	7	1.6	218	2 09RFT4	09RFT4 streptococ
27	7	1.6	219	5 017594	017594 caenorhabdl
28	7	1.6	220	2 09XCK1	09XCK1 streptococ
29	7	1.6	220	5 036036	036036 spironucleu
30	7	1.6	221	12 065002	065002 alfalfa mos
31	7	1.6	227	5 09BME4	09BME4 glossina mo
32	7	1.6	233	6 09MZB0	09MZB0 ovis arles
33	7	1.6	236	13 09DPS8	09DPS8 notochonia
34	7	1.6	264	12 085260	085260 potato viru
35	7	1.6	265	9 038155	038155 bacterioph
36	7	1.6	267	9 099363	099363 bacterioph
37	7	1.6	267	12 085261	085261 potato viru
38	7	1.6	267	12 098628	098628 potato viru
39	7	1.6	267	12 086842	086842 potato viru
40	7	1.6	267	12 099026	099026 potato viru
41	7	1.6	267	12 005476	005476 potato viru
42	7	1.6	267	12 009711	009711 potato viru
43	7	1.6	267	12 09WH17	09WH17 potato viru
44	7	1.6	267	12 09QNM9	09QNM9 potato viru
45	7	1.6	267	12 09DUL0	09DUL0 potato viru

ALIGNMENTS

RESULT 1
ID 09K07 PRELIMINARY; PRT; 382 AA.
AC 09K07;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PHOSPHOLIPASE A1, PUTATIVE.
GN NMB0464.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxId=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J., Nelson W.C., Gwinn M.L., Deboy R., Peterson J.D., Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A., Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B., Ralston M.D., Uitterlinden T.R., Khouri H., Qin H., Yamashiro J., Gill J., Scarlato V., Masignani V., Piva M., Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappelli R., Venter J.C.;
RA "Complete genome sequence of Neisseria meningitidis serogroup B strain MC58".
RT Science 287:1809-1815(2000).
RL EMBL: AE002403; AAF40901.1; -;
DR TIGR: NMB0464; -;
DR InterPro: IPR003187; PLAI.
DR Pfam: PF02253; PLAI; 1.
KW COMPLETE proteome.
SQ SEQUENCE 382 AA; 42714 MW; B468A802F062E836 CRC64;

Query Match 2.0%; Score 9; DB 2; Length 382;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 279 PEIRLTPV 287
|||
Db 221 PEIRLTPV 229

```

RESULT 2
ID 09J721 PRELIMINARY; PRT: 409 AA.
AC 09J721:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PUTATIVE PHOPHOLIPASE.
GN NMA2021.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=63699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=2022556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Felwell T., Hamlin N., Holroyd S.,
RA Jagers K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491."
RL Nature 404:502-506(2000).
DR EMBL; AL162757; CAB85240.1;
DR InterPro: IPR003187; Pfam:
DR Pfam: PR02253; Pfam: 1.
KW Complete proteome.
SQ SEQUENCE 409 AA; 45862 MW; CD6585B064D01A41 CRC64;

Query Match
Best Local Similarity 100.0%; Score 9; DB 2; Length 409;
Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIFLTQPV 287
DB 248 PEIFLTQPV 256

RESULT 3
ID 09CL62 PRELIMINARY; PRT: 862 AA.
AC 09CL62:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CVGSY.
GN CVGSY OR PM1380.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang O., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida PM70."
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
CC REGULATORY COMPONENTS OF SENSORY TRANSDUCTION SYSTEMS.
CC -1- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
CC KINASES.
DR EMBL; AE006176; AA033464.1;
DR InterPro: IPR000410; BcTL_sensory.
DR InterPro: IPR000658; DUF5.
DR InterPro: IPR003660; HAMP.
DR InterPro: IPR003594; HATPase_c.
DR InterPro: IPR003661; His_kinA.

```

```

DR InterPro: IPR001789; Response-reg.
DR Pfam: PF00672; DUF5. 1.
DR Pfam: PF00072; response-reg. 1.
DR Pfam: PF00512; signal. 1.
DR PRINTS: PR00344; BCTRSENSOR.
DR SMART: SM00304; HAMP. 1.
DR SMART: SM00387; HATPase_c. 1.
DR SMART: SM00388; HSKA. 1.
DR SMART: SM00448; RBC. 1.
KW Complete proteome; Kinase; Phosphorylation; Sensory transduction;
KW Transferrase.
SQ SEQUENCE 862 AA; 97846 MW; AB6814A3B63626C0 CRC64;

Query Match
Best Local Similarity 100.0%; Score 9; DB 2; Length 862;
Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LSTFLSTIL 13
DB 293 LSTFLSTIL 301

RESULT 4
ID 09AX85 PRELIMINARY; PRT: 1432 AA.
AC 09AX85:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PUTATIVE ABC TRANSPORTER PROTEIN.
GN P0410E03.7.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriaristaceae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone: P0410E03."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002844; BAB21276.1;
SQ SEQUENCE 1432 AA; 160234 MW; 04E8CF2CA08D993 CRC64;

Query Match
Best Local Similarity 100.0%; Score 9; DB 10; Length 1432;
Pred. No. 4.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 FTPLSLFD 178
DB 830 FTPLSLFD 838

RESULT 5
ID 09M268 PRELIMINARY; PRT: 299 AA.
AC 09M268:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE RAV-LIKE PROTEIN.
GN F21F14.140.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.

```

RA Choisme N., Robert C., Brothier P., Wincker P., Catolico L.,
RA Artiguenave F., Saurin W., Weissensbach J., Mewes H.W., Lemcke K.,
RA Mayer K.F.X., Queller F., Salanoubat M.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL38642; CAB/1904.1; -;
DR InterPro: IPR003340; B3.
DR Pfam: PF02362; B3; 1.
SQ SEQUENCE 299 AA; 34270 MW; 2AB841F179DD174B CRC64;

Query Match 1.8%; Score 8; DB 10; Length 299;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 SLSPDLR 181
DB 292 SLSPDLR 299

RESULT 6
O9CL22 PRELIMINARY; PRT; 306 AA.
AC O9CL22;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE HYPOTHETICAL PROTEIN PM1426.
GN PM1426.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PM70;
RX MEDLINE-21145866; PubMed-11248100;
RA May B.J., Zhang Q., Li L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL: AE006179; AAK03510.1; -;
DR InterPro: IPR003187; PLA1.
DR Pfam: PF02253; PLA1; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 306 AA; 35580 MW; EAF3DE8C1C22B26E CRC64;

Query Match 1.8%; Score 8; DB 2; Length 306;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 314 KLSRSMR 321
DB 184 KLSRSMR 191

RESULT 7
O67608 PRELIMINARY; PRT; 55 AA.
AC O67608;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE SUBGENOMIC DNA DERIVED FROM DNA B CCCDS - COVALENTLY CLOSED CIRCULAR
DE DOUBLE-STRANDED MOLECULE.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10831;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE-87040767; PubMed-3022243;
RA MacDowell S.W., Coutts R.H.A., Buck K.W.;
RT "Molecular characterization of subgenomic single-stranded and double-
RT stranded DNA forms isolated from plants infected with tomato golden
RT mosaic virus.";
RL Nucleic Acids Res. 14:7967-7984(1986).
DR EMBL: X04485; CAA28171.1; -;
DR InterPro: IPR000211; Gemin1.BL.
DR Pfam: PF00845; Gemin1.BL; 1.
SQ SEQUENCE 55 AA; 6281 MW; 41963570D739EC96 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 55;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 312 SAKLSRS 318
DB 43 SAKLSRS 49

RESULT 8
O9IZS5 PRELIMINARY; PRT; 72 AA.
ID O9IZS5;
AC O9IZS5;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y strain NTN (PVY (NTN)).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=122280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TUBER NECROSING (NTN);
RA Cerovska N., Filigaro M., Moravec T., Petrzik K.;
RT "Differences in nucleotide and amino acid sequences of N-terminal
RT parts of coat proteins among isolates of potato virus Y NTN strain.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF228633; AAF67874.1; -;
FT NON_TER 1 72
FT NON_TER 1 72
SQ SEQUENCE 72 AA; 7721 MW; 2C5287E15624CBDD CRC64;

Query Match 1.6%; Score 7; DB 12; Length 72;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 156 TKDKAQ 162
DB 11 TKDKAQ 17

RESULT 9
O9IZS4 PRELIMINARY; PRT; 72 AA.
ID O9IZS4;
AC O9IZS4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y strain NTN (PVY (NTN)).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=122280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TUBER NECROSING (NTN);
RA Cerovska N., Filigaro M., Moravec T., Petrzik K.;

RT "Differences in nucleotide and amino acid sequences of N-terminal
 parts of coat proteins among isolates of potato virus Y-NTN strain."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF228634; AAF67875.1; -.
 FT NON_TER 1 1
 SEQUENCE 72 AA; 7692 MW; 4F129B870523B482 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 72;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKQ 162
 Db 11 TKKDAKQ 17

RESULT 10
 ID Q91ZS3 PRELIMINARY; PRT; 95 AA.
 AC Q91ZS3;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DE COAT PROTEIN (FRAGMENT).
 GN CP.
 OS Potato virus Y strain NTN (PVY(NTN)).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 OC Polyvirus.
 OX NCBI_TaxID=122280;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-TUBER NECROSING (NTN);
 RA Cerovska N., Filligiarova M., Moravec T., Petrizik K.;
 RT "Differences in nucleotide and amino acid sequences of N-terminal
 parts of coat proteins among isolates of potato virus Y-NTN strain."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF228635; AAF67876.1; -.
 FT NON_TER 1 1
 SEQUENCE 95 AA; 10446 MW; 28C06891777C928 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 95;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKQ 162
 Db 11 TKKDAKQ 17

RESULT 11
 ID Q91ZS8 PRELIMINARY; PRT; 96 AA.
 AC Q91ZS8;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DE COAT PROTEIN (FRAGMENT).
 GN CP.
 OS Potato virus Y (strain N) (PVY).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 OC Polyvirus.
 OX NCBI_TaxID=12219;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-NECROSING (N);
 RA Cerovska N., Filligiarova M., Moravec T., Petrizik K.;
 RT "Differences in nucleotide and amino acid sequences of N-terminal
 parts of coat proteins among isolates of potato virus Y-NTN strain."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF228630; AAF67871.1; -.
 FT NON_TER 1 1
 SEQUENCE 96 AA; 10559 MW; 89C3A5675B9784CA CRC64;

Query Match 1.6%; Score 7; DB 12; Length 96;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKQ 162
 Db 11 TKKDAKQ 17

RESULT 12
 ID Q91ZS6 PRELIMINARY; PRT; 101 AA.
 AC Q91ZS6;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DE COAT PROTEIN (FRAGMENT).
 GN CP.
 OS Potato virus Y strain NTN (PVY(NTN)).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 OC Polyvirus.
 OX NCBI_TaxID=122280;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-TUBER NECROSING (NTN);
 RA Cerovska N., Filligiarova M., Moravec T., Petrizik K.;
 RT "Differences in nucleotide and amino acid sequences of N-terminal
 parts of coat proteins among isolates of potato virus Y-NTN strain."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF228632; AAF67873.1; -.
 FT NON_TER 1 1
 SEQUENCE 101 AA; 11112 MW; B5B5F51C3D5FE9BD CRC64;

Query Match 1.6%; Score 7; DB 12; Length 101;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKQ 162
 Db 11 TKKDAKQ 17

RESULT 13
 ID Q9RU16 PRELIMINARY; PRT; 122 AA.
 AC Q9RU16;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DE HYPOTHETICAL 13.7 KDA PROTEIN.
 GN DR1580.
 OS Deinococcus radiodurans.
 OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
 OX NCBI_TaxID=1299;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-RI;
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Milton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;

RT "Genome sequence of the radioresistant bacterium Deinococcus
 RT radiodurans RI.";
 RL Science 286:1571-1577(1999).
 DR EMBL: AE002002; AAF1150.1; -.
 DR TIGR: DR1580; -.
 DR InterPro: IPR000325; Glyoxalase_1.
 DR Pfam: PF00903; Glyoxalase: 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 122 AA; 13683 MW; 7BA836DCAD786E51 CRC64;

Query Match 1.6%; Score 7; DB 2; Length 122;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 EHEPELY 76
 |||||
 Db 115 EHEPELY 121

RESULT 14
 Q91ZS7 PRELIMINARY; PRT; 124 AA.
 AC Q91ZS7;
 DT 01-OCT-2000 (TREMUREL. 15, Created)
 DT 01-OCT-2000 (TREMUREL. 15, Last sequence update)
 DT 01-OCT-2000 (TREMUREL. 15, Last annotation update)
 DE COAT PROTEIN (FRAGMENT).
 GN CP.
 OS Potato virus Y strain NTN (PVY(NTN)).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 OC Polyviruses.
 OX NCBI_TaxID-122280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TUBER NECROSING (NTN):
 RA Cerovsky N., Filigayova M., Moravec T., Petrzik K.;
 RT "Differences in nucleotide and amino acid sequences of N-terminal
 parts of coat proteins among isolates of potato virus Y-NTN strain."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF228631; AAF67872.1; -.
 FT NON TER 1
 FT NON TER 124
 SQ SEQUENCE 124 AA; 13654 MW; DE3618D620BA7080 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 124;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAQ 162
 |||||
 Db 11 TKKDAQ 17

RESULT 15
 Q9GMD4 PRELIMINARY; PRT; 131 AA.
 AC Q9GMD4;
 DT 01-MAR-2001 (TREMUREL. 16, Created)
 DT 01-MAR-2001 (TREMUREL. 16, Last sequence update)
 DT 01-JUN-2001 (TREMUREL. 17, Last annotation update)
 DE ALPHA-TUBULIN (FRAGMENT).
 OS Macaca fuscata (Japanese macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID-9542;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99270338; PubMed=10340513;
 RA Okuno H., Tokuyama W., Li Y.X., Hashimoto T., Miyashita Y.;
 RT "Quantitative evaluation of neurotrophin and trk mRNA expression in

RT visual and limbic areas along the occipito-temporo-hippocampal pathway
 RT in adult macaque monkeys.";
 RL J. Comp. Neurol. 408:378-398(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Hashimoto T., Tokuyama W., Okuno H., Li Y., Miyashita Y.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF286378; AAC00533.1; -.
 DR InterPro: IPR003008; Tubulin_FtsZ.
 DR Pfam: PF00091; tubulin; 1.
 KW GTP-binding.
 FT NON TER 1
 FT NON TER 131
 SQ SEQUENCE 131 AA; 14575 MW; A07D9C90B0641FCE CRC64;

Query Match 1.6%; Score 7; DB 6; Length 131;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 IKTKRSI 120
 |||||
 Db 110 IKTKRSI 116

Search completed: May 9, 2002, 13:05:19
 Job time: 614 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 13:05:53 ; Search time 28.06 Seconds
(without alignments) 577.543 Million cell updates/sec

Title: US-09-787-083-8
Perfect score: 442
Sequence: 1 MKVSLSTLTSLPCFALTA.....YNHEATSPGVGLINLMDMGL 442

Scoring table: OLIGO
Gapop 60.0 , Gapept 60.0

Searched: 100059 seqs, 36664827 residues

Word size : 0

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	1.8	286	1	P37446 KLEPN
2	8	1.8	1396	1	VLTF_BP75
3	7	1.6	186	1	PRL_ICTPU
4	7	1.6	197	1	TBA_PLAYO
5	7	1.6	240	1	TBA_OCTVU
6	7	1.6	270	1	Y304_MYCPN
7	7	1.6	273	1	ILVE_THBMA
8	7	1.6	322	1	MIAA_PSEPU
9	7	1.6	326	1	SPL_BACBR
10	7	1.6	342	1	FLAD_METJA
11	7	1.6	342	1	Y735_METJA
12	7	1.6	423	1	AMT_AQUAE
13	7	1.6	444	1	TBA_ONCKE
14	7	1.6	448	1	TBA4_HUMAN
15	7	1.6	448	1	TBA5_CHICK
16	7	1.6	449	1	TBA2_DROME
17	7	1.6	449	1	TBAE_PHYPO
18	7	1.6	449	1	TBAN_PHYPO
19	7	1.6	449	1	TBA_XENLA
20	7	1.6	450	1	TBAD_PHYPO
21	7	1.6	450	1	TBA_HARCO
22	7	1.6	450	1	TBA_MORVI
23	7	1.6	451	1	TBA1_CRIGR
24	7	1.6	451	1	TBA3_HOMAM
25	7	1.6	451	1	TBA_OCTDO
26	7	1.6	451	1	TBA_TORMA
27	7	1.6	511	1	LGT_CITUN
28	7	1.6	513	1	ATPA_HAEIN
29	7	1.6	536	1	60IM_BUCAP
30	7	1.6	544	1	GPV_RAT
31	7	1.6	567	1	GPV_RAT
32	7	1.6	610	1	FIMB_DICDI
33	7	1.6	635	1	DNAM_HALMA

34	7	1.6	768	1	YB23_HUMAN	Q9U1J7	homo sapien
35	7	1.6	783	1	FXB_HUMAN	O151I7	homo sapien
36	7	1.6	900	1	AXN1_HUMAN	O151I6	homo sapien
37	7	1.6	947	1	LKTA_PASBP	P55123	pasteurella
38	7	1.6	1026	1	VG37_BP74	P03744	bacterioph
39	7	1.6	3061	1	POLG_PYHU	Q02963	p genome po
40	6	1.4	33	1	BR2B_RANES	P40838	rana escul
41	6	1.4	33	1	BR2E_RANES	P32413	rana escul
42	6	1.4	50	1	YK05_CABEL	P34300	caenorhabd
43	6	1.4	51	1	INS_ACOCA	P01324	acomys cahl
44	6	1.4	51	1	INS_ALIMI	P12703	alligator m
45	6	1.4	51	1	INS_ANSAN	P07454	anser anser

ALIGNMENTS

RESULT	1	STANDARD	PRT	286 AA
PAL_KLEPN				
AC P37446;				
DT 01-OCT-1994 (Rel. 30, Created)				
DT 01-OCT-1994 (Rel. 30, Last sequence update)				
DT 01-NOV-1997 (Rel. 35, Last annotation update)				
DE PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DEMERGENT-RESISTANT				
DE PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-				
DE ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).				
GN PLDA.				
OS Klebsiella pneumoniae.				
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;				
OC Klebsiella.				
OX NCBI_TaxID=573;				
RN [1]				
RN SEQUENCE FROM N.A.				
RP MEDLINE=94131966; PubMed=8300539;				
RA Brok R.G.P.M., Brinkman E., van Boxtel R., Bekkers A.C.A.P.,				
RA Verheij H.M., Tomassen J.;				
RT Molecular characterization of enterobacterial plda genes encoding				
RT outer membrane phospholipase A.;				
RL J. Bacteriol. 176:861-870(1994).				
CC A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.				
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O -> 1-ACYLGLYCERO-				
CC PHOSPHOCHOLINE + A FATTY ACID ANION.				
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O -> 2-ACYLGLYCERO-				
CC PHOSPHOCHOLINE + A FATTY ACID ANION.				
CC -1- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.				
CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES				
CC LOCATED THERE.				
CC				
CC This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC the European Bioinformatics Institute. There are no restrictions on its				
CC use by non-profit institutions as long as its content is in no way				
CC modified and this statement is not removed. Usage by and for commercial				
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/				
CC or send an email to license@isb-sib.ch).				
CC				
CC EMBL; X76901; CAA54223.1; -				
DR PIR; B36971; B36971.				
DR PIR; S40129; S40129.				
DR InterPro; IPR003187; PAL1.				
DR Pfam; PF02253; PLA1; 1.				
KW Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium.				
FT STGNAL	1	20		
FT CHAIN	21	286		
FT ACT SITE	161	161		
FT				
SQ	SEQUENCE	286 AA; 32544 MW; 3E39F863085108A3 CRC64;		

Query Match 1.8%; Score 8; DB 1; Length 286;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 431 GVGMLND 438
 DB 277 GVGMLND 284

RESULT 2

VFLE_BPT5 STANDARD: PRT: 1396 AA.
 AC P13390; 048502;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE L-SHAPED TAIL FIBER PROTEIN (LTF PROTEIN).
 GN LTF.
 OS Bacteriophage T5.
 OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae.
 OX NCBI_TaxID=10726;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95309401; PubMed=7789514;
 RA Kallman A.V., Kulshln V.E., Shlyapnikov M.G., Ksenzenko V.N.,
 RA Kiyukov V.M.;
 RT "The nucleotide sequence of the bacteriophage T5 ltf gene.";
 RL FEBS Lett. 366:46-48(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Kallman A.V.;
 RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
 RX MEDLINE=88289370; PubMed=3267228;
 RA Kallman A.V., Kiyukov V.M., Bayev A.A.;
 RT "The nucleotide sequence of bacteriophage T5 DNA at the region
 RT between early and late genes.";
 RL Nucleic Acids Res. 16:6230-6230(1988).
 CC -1- FUNCTION: NONESSENTIAL PROTEIN THAT MEDIATES BINDING TO THE
 CC POLYMANNOSE O ANTIGEN.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL: X69460; CAA49220.1; -;
 DR EMBL: AJ001191; CAA04591.1; -;
 DR PIR: S01982; S01982.

KW Late protein.
 FT CONFICIT 986 986 V -> A (IN REF. 2).
 SQ SEQUENCE 1396 AA; 147989 MW; 18CD2192F65FFC1 CRC64;

Query Match 1.8%; Score 8; DB 1; Length 1396;
 Best Local Similarity 100.0%; Pred. No. 6.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 54 QSATOSAS 61
 DB 87 QSATOSAS 94

RESULT 3
 PRL_ICTPU STANDARD: PRT: 186 AA.

AC P51504;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE PROLACTIN (PRL).
 OS Ictalurus punctatus (Channel catfish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Siluriformes; Ictaluridae; Ictalurus.
 OX NCBI_TaxID=7998;

RESULT 4

RP SEQUENCE.
 RC TISSUE=pituitary;
 RX MEDLINE=93364578; PubMed=1308206;
 RA Watanabe K., Igarashi A., Noso T., Chen T.T., Dunham R.A.,
 RA Kawachi H.;
 RT "Chemical identification of catfish growth hormone and prolactin.";
 RL Mol. Mar. Biol. Biotechnol. 1:239-249(1992).
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
 DR HSP: Q28632; IAN3.
 DR InterPro: IPR001400; SOMATOTROPIN.
 DR Pfam: PF00103; hormone; 1.
 DR PRINTS: PR00836; SOMATOTROPIN.
 DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
 DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
 KW Hormone; Pituitary.
 FT DISULFID 45 159 BY SIMILARITY.
 FT DISULFID 176 186 BY SIMILARITY.
 SQ SEQUENCE 186 AA; 20606 MW; 7BCD89718F44E74 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 186;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 41 NDIGDN 47
 DB 140 NDIGDN 146

RESULT 4
 TBA_PLAYO STANDARD: PRT: 197 AA.

AC P12543;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE TUBULIN ALPHA CHAIN (FRAGMENT).
 OS Plasmodium berghei yoelli.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5862;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89014607; PubMed=2459618;
 RA Akella R., Arasu P., Vaidya A.B.;
 RT "Molecular clones of alpha-tubulin genes of Plasmodium yoelli reveal
 RT an unusual feature of the carboxy terminus.";
 RL Mol. Biochem. Parasitol. 30:165-174(1988).
 CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
 CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
 CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
 CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL: M29816; AAA29779.1; -;
 DR PIR: A54506; A54506.
 DR InterPro: IPR000217; Tubulin.
 DR InterPro: IPR003008; Tubulin_ftsz.
 DR Pfam: PF00091; tubulin; 1.
 DR PROSITE: PS00227; TUBULIN; PARTIAL.

KW Microtubules: GTP-binding.
 FT NON_TER 1
 SQ SEQUENCE 197 AA; 22111 MW; 4DAAP199CC6CD319 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 197;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTKRSI 120
 |||||
 DB 82 IKTKRSI 88

RESULT 5
 TBA_OCTVU STANDARD; PRT; 240 AA.
 AC P2A635;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE TUBULIN ALPHA CHAIN (FRAGMENT).
 OS Octopus vulgaris (Octopus).
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda.
 CC Invertebrata; Octopodidae; Octopus.
 OX NCBI_TaxID=6645;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lens;
 RA Zinov'Eva R.D., Aleinikova K.S., Tomarev S.I.;
 RT "Isolation and structural characterization of cDNAs coding for alpha-tubulin of the octopus eye-lens."
 RL Dokl. Akad. Nauk SSSR 302:462-467(1988).
 CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X15845; CAA33844.1; -.
 DR PIR: A61544; A61544.
 DR InterPro: IPR000217; Tubulin.
 DR InterPro: IPR003008; Tubulin_FtsZ.
 DR Pfam: PF00091; Tubulin_1.
 DR PROSITE: PS00227; TUBULIN; PARTIAL.
 KW Microtubules; GTP-binding.
 FT NON_TER 1
 FT SITE 1
 SQ SEQUENCE 240 AA; 26961 MW; 13BB3A1F740E2416 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 240;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTKRSI 120
 |||||
 DB 124 IKTKRSI 130

RESULT 6
 Y304_MYCPN STANDARD; PRT; 270 AA.
 ID Y304_MYCPN
 AC P75355;
 DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOTHEICAL ABC TRANSPORTER ATP-BINDING PROTEIN MG304 HOMOLOG (A05_0RF270L).
 DE (A05_0RF270L).
 GN MPN433 OR MP408.
 OS Mycoplasma pneumoniae.
 CC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 CC Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;
 RA MEDLINE=97105885; PubMed=8948633;
 RA Himmelfreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C., Herrmann R.;
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae."
 RL Nucleic Acids Res. 24:4420-4449(1996).
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE000040; AAB96056.1; -.
 DR InterPro: IPR003593; AAA.
 DR InterPro: IPR003439; ABC_transportr.
 DR InterPro: IPR001687; ATP_GTP_A.
 DR Pfam: PF00005; ABC_tran; 1.
 DR SMART: SM00382; AAA; 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 KW Hypothetical protein; ATP-binding; Transport; Complete proteome.
 FT NP_BIND 36
 FT SITE 43
 SQ SEQUENCE 270 AA; 30770 MW; 6512640E4BC051B4 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 270;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 168 KQTPPLS 174
 |||||
 DB 125 KQTPPLS 131

RESULT 7
 ILVE_THEMA STANDARD; PRT; 273 AA.
 ID ILVE_THEMA
 AC P74921;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PROBABLE BRANCHED-CHAIN AMINO ACID AMINOTRANSFERASE (EC 2.6.1.42) (BCAT).
 DE (BCAT).
 GN ILVE OR TM0831.
 OS Thermotoga maritima.
 CC Bacteria; Thermotogales; Thermotoga.
 OX NCBI_TaxID=2336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RX MEDLINE=99287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J., Hart D.R., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A., McDonald L., Utterback T.R., Malek J.A., Liner K.D., Garrett M.M., Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D., Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O., Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;

RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
RN (2)
RP SEQUENCE OF 68-273 FROM N.A.
RX STRAIN-MSB8 / DSM 3109;
RC MEDLINE=97017137; PubMed=8863738;
RA Gulpud O., Labeledan B., Forrester P.;
RT "A gyrB-like gene from the hyperthermophilic bacterium Thermotoga
maritima.";
RL Gene 174:121-128(1996).
CC -1- CATALYTIC ACTIVITY: L-LEUCINE + 2-OXOGLOUTARATE -> 4-METHYL-2-
CC OXOBUTANOATE + L-GLUTAMATE (ALSO ACTS ON L-ISOLEUCINE AND
CC L-VALINE).
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: VALINE AND ISOLEUCINE BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO CLASS-IV OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE001750; AAD35913.1; -
DR EMBL: U49692; AAC44497.1; -
DR HSSP: P00510; 1A3G.
DR TIGR: TM0831; -
DR InterPro: IPR001544; Aminoctran_4.
DR Pfam: PF01063; aminotran_4; 1.
DR ProDom: PD001961; Aminoctran_4; 1.
DR PROSITE: PS00770; AA_TRANSFER_CLASS_4; 1.
KM Transferase: Aminoctrantransferase: Branched-chain amino acid biosynthesis;
FM Pyridoxal phosphate: Complete proteome.
RT CONFLICT 77
SQ SEQUENCE 273 AA; 31158 MW; 2162B705612E90E3 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 273;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 172 PLSLSPD 178
Db 61 PLSLSPD 67
|||||

RESULT 8
MIAA_PSEPU STANDARD; PRT; 322 AA.
AC 030762;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE TRNA DELTA(2)-ISOPENTENYL-PYROPHOSPHATE TRANSFERASE (EC 2.5.1.8) (IPP
DE TRANSFERASE) (ISOPENTENYL-DIPHOSPHATE:TRNA ISOPENTENYLTRANSFERASE)
DE (1PTASE) (1PTP).
GN MIAA.
OS Pseudomonas putida.
OC Bacteria: Proteobacteria: gamma subdivision: Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=503;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M.
RA Olekhnovich I.N., Gusein G.N.;
RT "Attenuation of the Pseudomonas putida trpE and trpGDC genes.";
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE FIRST STEP IN THE BIOSYNTHESIS OF
CC 2-METHYLTHIO-N6-(DELTA(2)-ISOPENTENYL)-ADENOSINE (MS[2]1[6]A1)

CC ADJACENT TO THE ANTICODON OF SEVERAL TRNA SPECIES (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ISOPENTENYL DIPHOSPHATE + TRNA -
CC PYROPHOSPHATE + TRNA CONTAINING 6-ISOPENTENYLADENOSINE.
CC -1- SIMILARITY: BELONGS TO THE IPP TRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaborati
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercia
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF016312; AAB69443.1; -
DR InterPro: IPR002627; IPPP.
DR Pfam: PF01715; IPPP; 1.
DR ProDom: PD004674; IPPP; 1.
KM Transferase: Nucleotidyltransferase: trna processing; ATP-binding.
FT NP_BIND 12 19
SQ SEQUENCE 322 AA; 35488 MW; DIA8B1A2D7B32F6A CRC64;

Query Match 1.6%; Score 7; DB 1; Length 322;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 FVDEVRS 39
Db 231 FVDEVRS 237
|||||

RESULT 9
SPI_BACBR STANDARD; PRT; 326 AA.
ID SPI_BACBR
AC P43131;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PROTEASE INHIBITOR PRECURSOR (BBRP1).
OS Bacillus brevis.
OC Bacteria: Firmicutes: Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Brevibacillus.
RX NCBI_TaxID=1393;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-32; 104-112 AND 122-136.
RC STRAIN=HPD31;
RX MEDLINE=92304060; PubMed=1610177;
RA Shiga Y., Hasegawa K., Tsuboi A., Yamagata H., Ueda S.;
RT "Characterization of an extracellular protease inhibitor of Bacillus
brevis HPD31 and nucleotide sequence of the corresponding gene.";
RL Appl. Environ. Microbiol. 58:525-531(1992).
CC -1- FUNCTION: SHOWS INHIBITORY ACTIVITY TOWARDS SERINE PROTEASES, SUCH
CC AS TRYPSIN, CHYMOTRYPSIN, AND SUBTILISIN. MAY FORM A TRYPSIN-
CC INHIBITOR COMPLEX IN A MOLAR RATIO OF 1:1. IT IS HEAT RESISTANT AT
CC NEUTRAL AND ACIDIC pHs.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- PTM: MAY UNDERGO ACTIVATION AFTER SECRETION. IT IS PRODUCED
CC EXTRACELLULARLY IN MULTIPLE FORMS HAVING AT LEAST THREE DIFFERENT
CC MOLECULAR WEIGHTS (BBRP1-A, -B, AND -C).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaborati
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercia
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D10696; BAA01538.1; -
DR Serine protease inhibitor: Protease inhibitor; Signal; Repeat.
KM SIGNAL 1 24
FT CHAIN 25 326
FT CHAIN 104 326
SERINE PROTEASE INHIBITOR.
SERINE PROTEASE INHIBITOR-C.

FT CHAIN 104 ? SERINE PROTEASE INHIBITOR-B (POTENTIAL).
 FT CHAIN 122 326 SERINE PROTEASE INHIBITOR-A.
 FT DOMAIN 177 304 CONTAINS TWO APPROXIMATE REPEATS.
 FT REPEAT 177 208 1.
 FT REPEAT 272 304 2.
 SQ SEQUENCE 326 AA; 35100 MM; 1C0456ABFA912E77 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 326;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 30 PVAFVDE 36
 Db 119 PVAFVDE 125

RESULT 10
 FLAD_METUA STANDARD: PRT: 342 AA.
 AC Q58305;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PUTATIVE FLAGELLA-RELATED PROTEIN D.
 GN FLAD OR M0895.
 OS Methanococcus jannaschli.
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
 OC Methanococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
 RA MEDLINE=96337999; PubMed=6688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschli.";
 RL Science 273:1058-1073(1996).
 CC -1- SIMILARITY: STRONG, TO M.VOLTAE FLAD, ALSO TO FLAE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U67533; AAB98898.1; -
 DR TIGR: M0755; -
 DR Flagella: Complete proteome.
 KW SEQUENCE 342 AA; 39950 MM; B384DDE1775566C CRC64;
 SQ

Query Match 1.6%; Score 7; DB 1; Length 342;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 356 PDILDY 362
 Db 255 PDILDY 261
 RESULT 11
 Y755_METUA STANDARD: PRT: 342 AA.
 ID Y755_METUA

AC 058165;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOTHETICAL PROTEIN M0755 PRECURSOR.
 GN M0755.
 OS Methanococcus jannaschli.
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
 OC Methanococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
 RA MEDLINE=96337999; PubMed=6688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschli.";
 RL Science 273:1058-1073(1996).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U67521; AAB98748.1; -
 DR TIGR: M0755; -
 DR Hypothetical protein: Signal: Complete proteome.
 KW Hypothetical protein: POTENTIAL.
 FT SIGNAL 1 18
 FT CHAIN 19 342 HYPOTHETICAL PROTEIN M0755.
 SQ SEQUENCE 342 AA; 38778 MM; BD25220A/EEND85C CRC64;

Query Match 1.6%; Score 7; DB 1; Length 342;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 76 YTTALEN 82
 Db 90 YTTALEN 96

RESULT 12
 AMT_AOUAE STANDARD: PRT: 423 AA.
 AC 066515;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PROBABLE AMMONIUM TRANSPORTER.
 GN AMT OR AMTB OR AO_112.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.
 OX NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VF5;
 RA MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aulay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RA "The complete genome of the hyperthermophilic bacterium Aquifex
 aeolicus.";
 RL Nature 392:353-358(1998).

```

CC -1- FUNCTION: INVOLVED IN THE UPTAKE OF AMONIA (PROBABLE).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE AMT1/MEP/MRGA FAMILY OF AMMONIUM
CC TRANSPORTERS (TC 2.4.9).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000674; AAC06478.1; -.
CC InterPro: IPR001905; Ammonium_transp.
CC Pfam: PF00909; Ammonium_transp. 1.
CC PROSITE: PS01219; AMMONIUM_TRANSPO. 1.
CC KW Transport; Transmembrane; Complete proteome.
CC FT TRANSMEM 3 23 POTENTIAL.
CC FT TRANSMEM 31 51 POTENTIAL.
CC FT TRANSMEM 64 84 POTENTIAL.
CC FT TRANSMEM 123 143 POTENTIAL.
CC FT TRANSMEM 150 170 POTENTIAL.
CC FT TRANSMEM 183 203 POTENTIAL.
CC FT TRANSMEM 217 227 POTENTIAL.
CC FT TRANSMEM 246 266 POTENTIAL.
CC FT TRANSMEM 278 298 POTENTIAL.
CC FT TRANSMEM 301 321 POTENTIAL.
CC FT TRANSMEM 331 351 POTENTIAL.
CC FT TRANSMEM 374 394 POTENTIAL.
CC SQ SEQUENCE 423 AA; 44322 MW; 166203BBB2A8A15B CRC64;

Query Match
Best Local Similarity 1.6%; Score 7; DB 1; Length 423;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 402 LKRGISG 408
DB 104 LKRGISG 110
IIIIIIII
TBA_ONCKE
ID TBA_ONCKE STANDARD; PRT; 444 AA.
AC P30436;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE TUBULIN ALPHA CHAIN.
OS Oncorhynchus keta (Chum salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8018;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93208376; PubMed=1296820;
RA Coe I.R., Munro R., Sherwood N.M.;
RT "Isolation of different brain-specific isoforms of alpha-tubulins
RT from chum salmon (Oncorhynchus keta).";
RL DNA Seq. 3:257-262(1992).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X66973; CAA47384.1; -.
CC PIR: S25004; S25004.
CC InterPro: IPR000217; Tubulin.
CC DR Pfam: PF003008; Tubulin_fcsz.
CC DR Pfam: PF00091; tubulin; 1.
CC DR PRINTS: PR01161; TUBULIN.
CC DR PRINTS: PR01162; ALPHATUBULIN.
CC PROSITE: PS00227; TUBULIN; 1.
CC KW Microtubules; GTP-binding; Multigene family.
CC FT NP_BIND 136 142 GTP (POTENTIAL).
CC FT SITE 444 444 INVOLVED IN POLYMERIZATION.
CC SQ SEQUENCE 444 AA; 49314 MW; C2CD75432DDDBAF CRC64;

Query Match
Best Local Similarity 1.6%; Score 7; DB 1; Length 444;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTKRSI 120
DB 329 IKTKRSI 335
IIIIIIII
TBA4_HUMAN
ID TBA4_HUMAN STANDARD; PRT; 448 AA.
AC P05215;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE TUBULIN ALPHA-4 CHAIN.
GN TUBA4.
OS Homo sapiens (Human).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606, 9541, 10090;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Human, and M.fascicularis;
RX MEDLINE=87146332; PubMed=3029670;
RA Dobner P.R., Kislauksis E., Wentworth B.M., Villa-Komaroff L.;
RT "Alternative 5' exons either provide or deny an initiator methionine
RT codon to the same alpha-tubulin coding region.";
RL Nucleic Acids Res. 15:199-218(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Mouse;
RX MEDLINE=87064538; PubMed=3785200;
RA Villasant A., Wang D., Dobner P., Dolph P., Lewis S.A., Cowan N.J.;
RT "Six mouse alpha-tubulin mRNAs encode five distinct isoforms: testis-
RT specific expression of two sister genes.";
RL Mol. Cell. Biol. 6:2409-2419(1986).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
CC -1- MISCELLANEOUS: THIS TUBULIN DOES NOT HAVE A C-TERMINAL TYROSINE.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

DR EMBL: X06956; CAA30026.1; -;
 DR EMBL: X04757; CAA28453.1; -;
 DR EMBL: M13444; AAA40502.1; -;
 DR PIR: A25873; A25873.
 DR MGD: MGI:1095410; Tubu4.
 DR InterPro: IPR000217; Tubulin.
 DR InterPro: IPR003008; Tubulin_FtsZ.
 DR Pfam: PF00091; tubulin; 1.
 DR PRINTS: PRO1161; TUBULIN.
 DR PRINTS: PRO1162; ALPHATUBULIN.
 DR PROSITE: PS00227; TUBULIN; 1.
 KM Microtubules; GTP-binding; Multigene family.
 FT NP_BIND 142 148 GTP (POTENTIAL).
 SQ SEQUENCE 448 AA; 49924 MW; C00ED90A183FE8F2 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 448;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTKRSI 120
 |||||
 Db 335 IKTKRSI 341

RESULT 15
 TBA5_CHICK STANDARD; PRT; 448 AA.
 AC P09644;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE TUBULIN ALPHA-5 CHAIN.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88296435; PubMed=3267229;
 RA Pratt L.F., Cleveland D.W.;
 RT "A survey of the alpha-tubulin gene family in chicken: unexpected
 RT sequence heterogeneity in the polypeptides encoded by five expressed
 RT genes.";
 RL EMBO J. 7:931-940(1988).
 CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
 CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
 CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
 CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
 CC -1- MISCELLANEOUS: THIS TUBULIN DOES NOT HAVE A C-TERMINAL TYROSINE.
 CC -1- MISCELLANEOUS: THERE ARE AT LEAST SEVEN ALPHA TUBULIN GENES
 CC (ALPHA-1 TO ALPHA-6, AND ALPHA-8), AND A PSEUDOGENE (ALPHA-7) IN
 CC CHICKEN.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X08061; CAA30852.1; ALT_INT.
 DR PIR: S00470; UBCHA5.
 DR InterPro: IPR000217; Tubulin.
 DR InterPro: IPR003008; Tubulin_FtsZ.
 DR Pfam: PF00091; tubulin; 1.
 DR PRINTS: PRO1161; TUBULIN.
 DR PRINTS: PRO1162; ALPHATUBULIN.
 DR PROSITE: PS00227; TUBULIN; 1.
 KM Microtubules; GTP-binding; Multigene family.

FT NP_BIND 142 148 GTP (POTENTIAL).
 SQ SEQUENCE 448 AA; 49948 MW; 2E498A2B20D7CBA6 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 448;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTKRSI 120
 |||||
 Db 335 IKTKRSI 341

Search completed: May 9, 2002, 13:05:54
 Job time: 609 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:56:34 ; Search time 44.37 Seconds

(without alignments)
758.827 Million cell updates/sec

Title: US-09-787-083-8

Perfect score: 442
Sequence: 1 MKVSLSTLTLSTLPCFAILA.....YNHEATSPFGVGLMDNMGL 442

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	2.0	382	2 E81195	phospholipase A1,
2	9	2.0	409	2 H81831	probable phospholip
3	8	1.8	286	2 B36971	outer membrane pho
4	8	1.8	299	2 T47989	RAV-like protein -
5	8	1.8	1396	2 S36851	L-shaped tail fibre
6	7	1.6	122	2 E75377	hypothetical prote
7	7	1.6	158	2 E86498	pts iia protein [i
8	7	1.6	158	2 E72124	tubulin alpha-1 ch
9	7	1.6	197	2 A54506	hypothetical prote
10	7	1.6	219	2 T19438	tubulin alpha chal
11	7	1.6	240	2 A61544	coat protein - pot
12	7	1.6	264	2 A44559	coat protein - pot
13	7	1.6	267	2 A60366	genome polypeptid
14	7	1.6	267	2 S14001	capsid protein - p
15	7	1.6	267	2 S26630	coat protein - pot
16	7	1.6	267	2 JCI527	tail fiber protein
17	7	1.6	267	2 S13339	abc transport ATP-
18	7	1.6	270	2 S73734	transaminase B hom
19	7	1.6	273	2 C73238	hypothetical phosph
20	7	1.6	288	2 T27985	hypothetical prote
21	7	1.6	306	2 T27985	probable RAV-like
22	7	1.6	310	2 T02200	polyprotein - pota
23	7	1.6	313	2 JT0960	conserved hypothet
24	7	1.6	321	2 H82240	hypothetical prote
25	7	1.6	324	2 T23876	protease inhibitor
26	7	1.6	336	2 A43939	hypothetical prote
27	7	1.6	338	2 F69437	pectate lyase-like
28	7	1.6	341	2 T47653	hypothetical prote
29	7	1.6	342	2 G64411	hypothetical prote

30	7	1.6	342	2 C64394	hypothetical prote
31	7	1.6	349	2 T43920	yfuc protein [impo
32	7	1.6	370	2 G83219	probable dihydroli
33	7	1.6	379	2 S13556	genome polypeptid
34	7	1.6	381	2 G82132	ribonuclease D VCI
35	7	1.6	382	2 S13237	tail fiber protein
36	7	1.6	384	2 S70638	tubulin alpha chal
37	7	1.6	384	2 S70641	tubulin alpha chal
38	7	1.6	384	2 S70640	tubulin alpha chal
39	7	1.6	387	2 S33517	tubulin alpha chal
40	7	1.6	407	2 T19895	hypothetical prote
41	7	1.6	415	2 H70477	UDP-N-acetylmutamo
42	7	1.6	419	2 F85064	hypothetical prote
43	7	1.6	419	2 T25577	hypothetical prote
44	7	1.6	423	1 UBFAA	tubulin alpha-1 ch
45	7	1.6	423	2 H70310	ammonium transport

ALIGNMENTS

RESULT 1

E81195 phospholipase A1, probable NMB0464 [imported] - Neisseria meningitidis (strain MC58 s

C:Species: Neisseria meningitidis

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C:Accession: E81195

R:Rettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.

ri, H.; Qian, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappoli, R.

A>Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755

A:Accession: E81195

A:Molecule type: DNA

A:Status: preliminary

A:Residues: 1-382 <TEXT>

A:Cross-references: GB:AE002403; GB:AE002098; NID:g7225688; PIDN:AAF40901.1; PID:g722

A:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: NMB0464

Query Match 2.0%; Score 9; DB 2; Length 382;

Best Local Similarity 100.0%; Pred.No. 0.45;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIFLTQPV 287

Db 221 PEIFLTQPV 229

RESULT 2

H81831 probable phospholipase NMA2021 [imported] - Neisseria meningitidis (strain Z2491 serog

C:Species: Neisseria meningitidis

C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C:Accession: H81831

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo

Nature 404, 502-506, 2000

A>Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491

A:Reference number: A81775; MUID:20222556

A:Accession: H81831

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-409 <PAR>

A:Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85240.1; PID:g738

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: NMA2021

Query Match 2.0%; Score 9; DB 2; Length 409;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEFLTPPV 287
DB 248 PEFLTPPV 256

RESULT 3
B36971
Outer membrane phospholipase A (EC 3.1.1.-) precursor - *Klebsiella pneumoniae*
C:Species: *Klebsiella pneumoniae*
C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 18-Jun-1999
C:Accession: B36971; S40129
R:Brok, R.G.P.M.; Brinkman, E.; van Boxel, R.; Bekkers, A.C.A.P.A.; Verheij, H.M.; Tomm
J. *Bacteriol.* 176, 861-870, 1994
A:Title: Molecular characterization of enterobacterial pldA genes encoding outer membran
A:Reference number: A36971; MUID:94131966
A:Accession: B36971
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-286 <BRO>
A:Cross-references: EMBL:X76901; NID:g436880; PIDN:CAA54223.1; PID:g436881
A:Note: authors translated the codon AAG for residue 112 as Arg
C:Genetics:
A:Gene: pldA
C:Superfamily: bacterial phospholipase A1
C:Keywords: carboxylic ester hydrolase

Query Match 1.8%; Score 8; DB 2; Length 286;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 431 GVGGLMND 438
DB 277 GVGGLMND 284

RESULT 4
T47989
RAY-like protein - *Arabidopsis thaliana*
N:Alternate names: protein F21F14.140
C:Species: *Arabidopsis thaliana* (mouse-ear cross)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47989
R:Sholsue, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Sa
submitted to the Protein Sequence Database, February 2000
A:Reference number: Z24481
A:Accession: T47989
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-299 <CHO>
A:Cross-references: EMBL:AL138642
A:Experimental source: cultivar Columbia; BAC clone F21F14
C:Genetics:
A:Map position: 3
A:Note: F21F14.140

Query Match 1.8%; Score 8; DB 2; Length 299;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 SLSFDDDR 181
DB 292 SLSFDDDR 299

RESULT 5
S36851

L-shaped tail fiber protein - phage T5
N:Alternate names: ltf protein
C:Species: phage T5
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 17-Nov-2000
C:Accession: S65934; S01984; S36851
R:Kallman, A.V.; Kulshin, V.E.; Shlyapnikov, M.G.; Ksenzenko, V.N.; Kryukov, V.M.
FEBS Lett. 366, 46-48, 1995
A:Title: The nucleotide sequence of the bacteriophage T5 ltf gene.
A:Reference number: S65934; MUID:95309401
A:Accession: S65934
A:Molecule type: DNA
A:Residues: 1-1396 <KAL>
A:Cross-references: EMBL:X69460; NID:g15415; PIDN:CAA49220.1; PID:g15416
R:Kallman, A.V.; Kryukov, V.M.; Bayev, A.A.
Nucleic Acids Res. 16, 6230, 1988
A:Title: The nucleotide sequence of bacteriophage T5 DNA at the region between early
A:Reference number: S01982; MUID:88289370
A:Accession: S01984
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 934-985, 'A', 987-1396 <KA2>
A:Cross-references: EMBL:X07559
C:Genetics:
A:Gene: ltf
C:Keywords: late protein; tail fiber

Query Match 1.8%; Score 8; DB 2; Length 1396;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 QSATQSAS 61
DB 87 QSATQSAS 94

RESULT 6
E75377
Hypothetical protein - *Deinococcus radiodurans* (strain R1)
C:Species: *Deinococcus radiodurans*
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: E75377
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
M.; Shen, M.; Yamathavan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
A:Reference number: A75250; MUID:20036896
A:Accession: E75377
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-122 <WHI>
A:Cross-references: GB:AE002002; GB:AE00513; NID:g6459345; PIDN:AAF1150.1; PID:g645
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRI580
A:Map position: 1

Query Match 1.6%; Score 7; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 EHEPELY 76
DB 115 EHEPELY 121

RESULT 7
E86498
Pta IIA Protein [imported] - *Chlamydomonas reinhardtii* (strain J138)
C:Species: *Chlamydomonas reinhardtii*
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001

C:Accession: E86498
 R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
 Nucleic Acids Res. 28, 2311-2314, 2000
 A:Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.
 A:Reference number: A86491; MUID:20330349
 A:Accession: E86498
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-158 <STO>
 A:Cross-references: GB:BA000008; NID:98978434; PIDN:BA098271.1; GSPDB:GN00142
 A:Experimental source: strain J138
 C:Genetics:
 A:Gene: ptnN_1
 C:superfamily: phosphotransferase system enzyme II; phosphotransferase system mannitol-5

Query Match 1.6%; Score 7; DB 2; Length 158;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LSTLTLS 11
 |||||
 DB 126 LSTLTLS 132

RESULT 8
 E72124
 p15 lla protein - Chlamydia pneumoniae (strain CWL029)
 C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
 C:Accession: E72124
 R:Rahman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
 Nature Genet. 21, 385-389, 1999
 A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
 A:Reference number: A72000; MUID:99206606
 A:Accession: E72124
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-158 <ARN>
 A:Cross-references: GB:AE001591; GB:AE001363; NID:94376311; PIDN:AAD18213.1; PID:9437631
 A:Experimental source: strain CWL029
 C:Genetics:
 A:Gene: ptnN_1
 C:superfamily: phosphotransferase system enzyme II; phosphotransferase system mannitol-5
 F:28-156/Domain: phosphotransferase system mannitol-specific enzyme II factor III homolog

Query Match 1.6%; Score 7; DB 2; Length 158;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LSTLTLS 11
 |||||
 DB 126 LSTLTLS 132

RESULT 9
 A54506
 tubulin alpha-1 chain - Plasmodium yoelii (fragment)
 C:Species: Plasmodium yoelii
 C:Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 13-Aug-1999
 C:Accession: A54506
 R:AKella, R.; Arasu, P.; Vaidya, A.B.
 Mol. Biochem. Parasitol. 30, 165-174, 1988
 A:Title: Molecular clones of alpha-tubulin genes of Plasmodium yoelii reveal an unusual
 A:Reference number: A54506; MUID:89014607
 A:Accession: A54506
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-197 <AKE>
 A:Cross-references: GB:M29816; NID:q160729; PIDN:AAA29779.1; PID:q160730
 C:superfamily: tubulin

Query Match 1.6%; Score 7; DB 2; Length 197;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 IKTKRSI 120
 |||||
 DB 82 IKTKRSI 88

RESULT 10
 T19438
 hypothetical protein C25A1.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T19438
 R:Mortimore, B.
 submitted to the EMBL Data Library, October 1996
 A:Reference number: Z19124
 A:Accession: T19438
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-219 <WIL>
 A:Cross-references: EMBL:Z81038; PIDN:CAB02762.1; GSPDB:GN00019; CESP:C25A1.1
 A:Experimental source: clone C25A1
 C:Genetics:
 A:Gene: CESP:C25A1.1
 A:Map position: 1
 A:introns: 33/1; 66/1; 151/2

Query Match 1.6%; Score 7; DB 2; Length 219;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 158 KDAKOLE 164
 |||||
 DB 52 KDAKOLE 58

RESULT 11
 A61544
 tubulin alpha chain - common octopus (fragment)
 C:Species: Octopus vulgaris (common octopus)
 C:Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 10-Jul-1998
 C:Accession: A61544
 R:Zinov'eva, R.D.; Aleinikova, K.S.; Tomarev, S.I.
 Dokl. Akad. Nauk SSSR 302, 462-467, 1988
 A:Title: Isolation and structural characterization of cDNAs coding for alpha-tubulin
 A:Reference number: A61544
 A:Accession: A61544
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-240 <ZIN>
 C:superfamily: tubulin
 F:239-240/Cleavage site: Glu-Tyr (tubulin-specific carboxypeptidase) #status predicte
 F:239-240/Cross-link: peptide (Glu-Tyr) (by tubulin-tyrosine ligase) #status predicte

Query Match 1.6%; Score 7; DB 2; Length 240;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 IKTKRSI 120
 |||||
 DB 124 IKTKRSI 130

RESULT 12
 A44959
 coat protein - potato virus Y (strain N) (fragment)
 C:Species: potato virus Y, PVY
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Nov-2000

C:Accession: A44959
R:Hay, J.M.; Fellowes, A.P.; Timmerman, G.M.
Arch. Virol. 107, 111-122, 1989
A:Title: Nucleotide sequence of the coat protein gene of a necrotic strain of potato virus Y
A:Reference number: A44959; MUID:90025790
A:Accession: A44959
A:Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-264 <HAY>
A:Cross-references: GB:M22470; NID:9333508; PIDN:AAA47182.1; PID:9333509
C:Superfamily: tobacco etch virus genome polypeptide

Query Match 1.6%; Score 7; DB 2; Length 264;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKQ 162
|||||||
DB 8 TKKDAKQ 14

RESULT 13

A60366
coat protein - potato virus Y (strain TH)
C:Species: potato virus Y, PVY
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Nov-2000
C:Accession: A60366
R:Hattaya, T.; Sano, T.; Ohshima, K.; Shikata, E.
Virus Genes 4, 339-350, 1990
A:Title: Polymerase chain-reaction-mediated cloning and expression of the coat protein gene
A:Reference number: A60366; MUID:91143125
A:Accession: A60366
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-267 <HAT>
C:Superfamily: tobacco etch virus genome polypeptide

Query Match 1.6%; Score 7; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKQ 162
|||||||
DB 11 TKKDAKQ 17

RESULT 14

S14001
genome polypeptide - potato virus Y (strain H) (fragment)
N:Contains: coat protein
C:Species: potato virus Y, PVY
A:Variety: strain H
C:Date: 21-Nov-1993 #sequence_revision 08-Nov-1996 #text_change 17-Nov-2000
C:Accession: S14001
R:Dalmay, T.; Balazs, E.
Nucleic Acids Res. 18, 6721, 1990
A:Title: Nucleotide sequence of an altered virulence potato virus Y coat protein gene (F)
A:Reference number: S14001; MUID:91067494
A:Accession: S14001
A:Status: translation not shown
A:Molecule type: genomic RNA
A:Residues: 1-267 <DAL>
A:Cross-references: EMBL:X54611; NID:961449; PIDN:CAA38432.1; PID:9930257
A:Experimental source: strain H
C:Superfamily: tobacco etch virus genome polypeptide
C:Keywords: coat protein; polypeptide
F:1-267/Product: coat protein #status predicted <CP>

Query Match 1.6%; Score 7; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 42;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 156 TKKDAKQ 162
|||||||
DB 11 TKKDAKQ 17

RESULT 15

S26630
capsid protein - potato virus Y (fragment)
C:Species: potato virus Y, PVY
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 17-Nov-2000
C:Accession: S26630
R:Sudarsono, X.Y.Z.; Woloshuk, S.L.; Lommel, S.A.; Xiong, Z.; Hellman, G.M.; Wernsman
submitted to the EMBL Data Library, September 1992
A:Description: Nucleotide sequence of the capsid protein cistrons from six potato virus
A:Reference number: S26628
A:Accession: S26630
A:Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-267 <SUD>
A:Cross-references: EMBL:X68223; NID:961435; PIDN:CAA48303.1; PID:961436
C:Superfamily: tobacco etch virus genome polypeptide

Query Match 1.6%; Score 7; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKQ 162
|||||||
DB 11 TKKDAKQ 17

Search completed: May 9, 2002, 12:56:35
Job time: 605 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 13:05:52 ; Search time 28.06 Seconds
(without alignments)
577.543 Million cell updates/sec

Title: US-09-787-083-6

Sequence: 442
1 MKVSLSTLTSLSCFAILA.....YNHEATSPFGVGLMDMNGL 442

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 100059 seqs, 36664827 residues

Word size : 0

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	1.8	286	1	PAL_KLEPN
2	8	1.8	1396	1	VITEF-BPT5
3	7	1.6	186	1	PRL_ICTPU
4	7	1.6	197	1	TBA_PLAYO
5	7	1.6	237	1	TRMD_BUCAI
6	7	1.6	240	1	TBA_OCTVU
7	7	1.6	270	1	Y304_MYCPN
8	7	1.6	273	1	ILVE_THEMA
9	7	1.6	322	1	MIAA_PSEPU
10	7	1.6	326	1	SPI_BACBR
11	7	1.6	342	1	FLAD_METJA
12	7	1.6	342	1	Y755_METJA
13	7	1.6	347	1	NU2M_CAPII
14	7	1.6	347	1	NU2M_HUMAN
15	7	1.6	347	1	NU2M_PANTR
16	7	1.6	347	1	NU2M_PONPA
17	7	1.6	347	1	NU2M_SHEEP
18	7	1.6	423	1	AMT_AQUAE
19	7	1.6	443	1	TBA_ONCKE
20	7	1.6	448	1	TBA4_HUMAN
21	7	1.6	448	1	TBA5_CHICK
22	7	1.6	449	1	TBA2_DROME
23	7	1.6	449	1	TBAE_PHYPO
24	7	1.6	449	1	TBAN_PHYPO
25	7	1.6	449	1	TBA_XENLA
26	7	1.6	450	1	TBAD_PHYPO
27	7	1.6	450	1	TBA_HARCO
28	7	1.6	450	1	TBA_NOTVI
29	7	1.6	451	1	TBA1_CRIGR
30	7	1.6	451	1	TBA3_HOMAM
31	7	1.6	451	1	TBA_OCTDO
32	7	1.6	451	1	TBA_TORMA
33	7	1.6	511	1	LGT_CITUN

34	7	1.6	513	1	ATPA_HAEIN	P43714 haemophilus
35	7	1.6	536	1	60IM_BUCAP	P29431 buchnera ap
36	7	1.6	544	1	GP10_DICDI	006885 dictyostell
37	7	1.6	567	1	GPV_RAT	008770 ratius norv
38	7	1.6	610	1	FIMB_DICDI	P54680 dictyostell
39	7	1.6	768	1	YB23_HUMAN	094117 homo sapien
40	7	1.6	783	1	FYB_HUMAN	015117 homo sapien
41	7	1.6	900	1	AXNI_HUMAN	015169 homo sapien
42	7	1.6	947	1	LKTA_PASSP	P55123 pasteurella
43	7	1.6	1026	1	VG37_BPT4	P03740 bacterioph
44	7	1.6	3061	1	POLG_PVYHU	002963 p genome po
45	6	1.4	33	1	BR2B_RANES	P40838 rana escul

ALIGNMENTS

RESULT	ID	STANDARD	PRT	286 AA
1	PAL_KLEPN			
AC	P37446			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	01-OCT-1994 (Rel. 30, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DEFERENT-RESISTANT			
DE	PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-			
DE	ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).			
GN	PLDA.			
OS	Klebsiella pneumoniae.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Klebsiella.			
ON	NCBI_TaxID=573;			
RX	SEQUENCE FROM N.A.			
RA	BROK R.G.P.M., Brinkman E., van Bortel R., Bekkers A.C.A.P.,			
RA	Verheij H.M., Tomassen J.,			
RT	outer membrane phospholipase A."			
RT	Molecular characterization of enterobacterial plda genes encoding			
RL	J. Bacteriol. 176:861-870(1994).			
CC	- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE			
CC	A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.			
CC	- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-			
CC	PHOSPHOCHOLINE + A FATTY ACID ANION.			
CC	- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-			
CC	PHOSPHOCHOLINE + A FATTY ACID ANION.			
CC	- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.			
CC	- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES			
CC	LOCATED THERE.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: X76901; CAAS4223.1; -			
DR	PIR: B36971; B36971.			
DR	PIR: S40129; S40129.			
DR	InterPro: IPR003187; PLA1.			
DR	Plam: PF02253; PLA1; 1.			
KW	Hydrolase; Lipid degradation; Outer membrane; signal; Calcium.			
FT	SIGNAL	1	20	BY SIMILARITY.
FT	CHAIN	21	286	PHOSPHOLIPASE A1.
FT	ACT_SITE	161	161	BY SIMILARITY.
FT	SEQUENCE	286 AA;	32544 MW;	3E39F863085108A3 CRC64;

Query Match 1.8%; Score 8; DB 1; Length 286;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 431 GVGGLMIND 438
 DB 277 GVGGLMIND 284

RESULT 2

VFLE_BPT5 STANDARD; PRT: 1396 AA.
 AC P13390; 048502;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE L-SHAPED TAIL FIBER PROTEIN (LTF PROTEIN).
 GN LTF.
 OS Bacteriophage T5.
 OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae.
 OX NCBI_TaxID=10726;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95309401; PubMed=7789514;
 RA Kallman A.V., Kulshin V.E., Shlyapnikov M.G., Ksenzenko V.N.,
 RA Krynukov V.M.;
 RT "The nucleotide sequence of the bacteriophage T5 ltf gene.";
 RT FEBS Lett. 366:46-48(1995).
 RL [2]
 RP SEQUENCE FROM N.A.
 RA Kallman A.V.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
 RX MEDLINE=88289370; PubMed=3267228;
 RA Kallman A.V., Krynukov V.M., Bayev A.A.;
 RT "The nucleotide sequence of bacteriophage T5 DNA at the region
 RT between early and late genes.";
 RL Nucleic Acids Res. 16:6230-6230(1988).
 CC -1- FUNCTION: NONESSENTIAL PROTEIN THAT MEDIATES BINDING TO THE
 CC POLYMANNOSE O ANTIGEN.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: X69460; CAA49220.1; -
 DR EMBL: AJ001191; CAA04591.1; -
 DR PIR: S01982; S01982.
 KW Late protein.
 FT CONFICT 986 V -> A (IN REF. 2).
 SQ SEQUENCE 1396 AA; 147989 MW; 18CD2192F65FFC1 CRC64;

Query Match 1.8%; Score 8; DB 1; Length 1396;
 Best Local Similarity 100.0%; Pred. No. 7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 54 QSATOSAS 61
 DB 87 QSATOSAS 94

RESULT 3

PRL_ICTPU STANDARD; PRT: 186 AA.
 AC P51904;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE PROLACTIN (PRL).
 OS Ictalurus punctatus (Channel catfish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Siluriformes; Ictaluridae; Ictalurus.
 OX NCBI_TaxID=7998;
 RN [1]
 RP SEQUENCE.

RC TISSUE-Pituitary;
 RX MEDLINE=93364578; PubMed=1308206;
 RA Watanabe K., Igarashi A., Noso T., Chen T.T., Dunham R.A.,
 RA Kawachi H.;
 RT "Chemical identification of catfish growth hormone and prolactin.";
 RL MOL. Mar. Biol. Biotechnol. 1:239-249(1992).
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
 DR HSP: Q28632; IAN3.
 DR InterPro: IPR001400; SOMATOTROPIN.
 DR Pfam: PF00103; hormone; 1.
 DR PRINTS: PR00836; SOMATOTROPIN.
 DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
 DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
 KW Hormone; Pituitary.
 FT DISULFID 45 159 BY SIMILARITY.
 FT DISULFID 176 186 BY SIMILARITY.
 SQ SEQUENCE 186 AA; 20606 MW; 7BCCD89718F44E74 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 186;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 41 NDIGDON 47
 DB 140 NDIGDON 146

RESULT 4

TBA_PLAYO STANDARD; PRT: 197 AA.
 AC P12543;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE TUBULIN ALPHA CHAIN (FRAGMENT).
 OS Plasmodium berghei yoellii.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5862;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89014607; PubMed=2459618;
 RA Akella R., Arasu P., Valodia A.B.;
 RT "Molecular clones of alpha-tubulin genes of Plasmodium yoellii reveal
 RT an unusual feature of the carboxy terminus.";
 RL MOL. Biochem. Parasitol. 30:165-174(1988).
 CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
 CC BINDS TWO MOLES OF GTP. ONE AT AN EXCHANGEABLE SITE ON THE BETA
 CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.

CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: M29816; AAA29779.1; -
 DR PIR: A54506; A54506.
 DR InterPro: IPR000217; Tubulin.
 DR InterPro: IPR003008; Tubulin_ftsz.
 DR Pfam: PF00091; tubulin; 1.
 DR PROSITE: PS00227; TUBULIN; PARTIAL.

KW Microtubules; GTP-binding.
 FT NON_TER 1
 SQ SEQUENCE 197 AA; 22111 MW; 4DAFA199CCGCD319 CRC64;

Query Match
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTKRSI 120
 |||||
 Db 82 IKTKRSI 88

RESULT 5
 TRMD_BUCAL STANDARD; PRT; 237 AA.

AC P57476;
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE tRNA (GUANINE-N1)-METHYLTRANSFERASE (EC 2.1.1.31) (MIG-
 DE METHYLTRANSFERASE) (tRNA [GM37] METHYLTRANSFERASE).
 GN TRMD OR BU396.
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
 symbiotic bacterium).
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
 OX NCBI_TaxID=118099;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TOKYO 1998;
 RX MEDLINE=20445173; PubMed=10993077;
 RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.,
 RT Buchnera sp. APS.";
 RT Genome sequence of the endocellular bacterial symbiont of aphids
 Nature 407:81-86(2000).
 RL Nature 407:81-86(2000).
 CC -1- FUNCTION: SPECIFICALLY METHYLATES GUANOSINE-37 IN VARIOUS TRNAS
 (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + TRNA - S-ADENOSYL-
 L-HOMOCYSTEINE + TRNA CONTAINING N(1)-METHYLGUANINE.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE RNA METHYLTRANSFERASE TRMD FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AP001119; BAB13099.1; -
 DR InterPro: IPR002649; tRNA_mig_MT.
 DR Pfam: PF01746; tRNA_mig_MT.1.
 DR ProDom: PD004978; tRNA_mig_MT.1.
 KW Transferase; Methyltransferase; tRNA processing; Complete proteome.
 SQ SEQUENCE 237 AA; 27335 MW; 86CE3836B35EC4AB CRC64;

Query Match
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LAIOQAK 25
 |||||
 Db 78 LAIOQAK 84

RESULT 6
 TBA_OCTVU STANDARD; PRT; 240 AA.
 AC P24635;
 DT 01-MAR-1992 (Rel. 21, Created)

DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE TUBULIN ALPHA CHAIN (FRAGMENT).
 OS Octopus vulgaris (Octopus).
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;
 OC Incirrata; Octopodidae; Octopus.
 OX NCBI_TaxID=6645;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue=Lens;
 RA Zinov'Eva R.D., Aleinikova K.S., Tomarev S.I.;
 RT "Isolation and structural characterization of cDNAs coding for alpha-
 RT tubulin of the octopus eye lens.";
 RL Dokl. Akad. Nauk SSSR 302:462-467(1988).
 CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
 CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
 CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
 CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X15845; CAA33844.1; -
 DR PIR: A61544; A61544.
 DR InterPro: IPR000217; Tubulin.
 DR InterPro: IPR003008; Tubulin_FtsZ.
 DR Pfam: PF00091; tubulin.1.
 DR PROSITE: PS00227; TUBULIN; PARTIAL.
 KW Microtubules; GTP-binding.
 FT NON_TER 1
 FT SITE 240 240 INVOLVED IN POLYMERIZATION.
 SQ SEQUENCE 240 AA; 26961 MW; 13BB3A1F740F2416 CRC64;

Query Match
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTKRSI 120
 |||||
 Db 124 IKTKRSI 130

RESULT 7
 Y304_MYCPN STANDARD; PRT; 270 AA.

AC P75355;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOTHEICAL ABC TRANSPORTER ATP-BINDING PROTEIN MG304 HOMOLOG
 DE (A05_ORF2701).
 GN MPN433 OR MP408.
 GN Mycoplasma pneumoniae.
 OS Mycoplasma pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molluscites;
 OC Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;
 RX MEDLINE=97105885; PubMed=8948633.
 RA Himmelfreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
 RT Herrmann R.;
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
 RT pneumoniae.";
 RL Nucleic Acids Res. 24:4420-4449(1996).
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY

(ABC TRANSPORTERS).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AEO00040; AAB96056.1; -
DR InterPro: IPR003593; AAA.
DR InterPro: IPR003439; ABC_transport.
DR Pfam: PF00005; ABC_tran.1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
KM Hypothetical protein: ATP-binding; Transport; Complete proteome.
FT NP_BIND 36 43 ATP (POTENTIAL).
SQ SEQUENCE 270 AA; 30770 MW; 6512640E4BC051B4 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 270;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 168 KQFTPLS 174
Db 125 KQFTPLS 131

RESULT 8
MIAA_PSEPU
ID ILVE_THEME STANDARD; PRT; 273 AA.
AC P74921;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PROBABLE BRANCHED-CHAIN AMINO ACID AMINOTRANSFERASE (EC 2.6.1.42)
DE (BCAA).
DE ILVE OR TM0831.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxId=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; Pubmed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heideberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima."
RL Nature 399:323-329(1999).
RN [2]
RP SEQUENCE OF 68-273 FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=97017137; Pubmed=8863738;
RA Gulpud O., Labeledan B., Forcierre P.;
RT "A gyrB-like gene from the hyperthermophilic bacterium Thermotoga
RT maritima."
RL Gene 174:121-128(1996).
CC -1- CATALYTIC ACTIVITY: L-LEUCINE + 2-OXOGLOUTARATE -> 4-METHYL-2-
CC OXOBUTANOATE + L-GLUTAMATE (ALSO ACTS ON L-ISOLEUCINE AND
CC L-VALINE).
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: VALINE AND ISOLEUCINE BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO CLASS-IV OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AEO01750; AAD35913.1; -
DR EMBL: U49692; AAC44497.1; -
DR HSSP: P00510; IACG.
DR TIGR: TM0831; -
DR InterPro: IPR001544; Aminotran.4.
DR Pfam: PF01063; aminotran_4; 1.
DR PRODOM: PD001961; Aminotran_4; 1.
DR PROSITE: PS00770; AA_TRANSFER_CLASS.4; 1.
KM Transferase: Aminotransferase; Branched-chain amino acid biosynthesis;
KM Pyridoxal phosphate; Complete proteome.
FT CONFLICT 77 A -> R (IN REF. 2).
SQ SEQUENCE 273 AA; 31158 MW; 2162B705612E90E3 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 273;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 172 PLSTSPD 178
Db 61 PLSTSPD 67

RESULT 9
MIAA_PSEPU
ID MIAA_PSEPU STANDARD; PRT; 322 AA.
AC Q30762;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE tRNA DELTA(2)-ISOPENTENYL-PYROPHOSPHATE TRANSFERASE (EC 2.5.1.8) (IPP
DE TRANSFERASE) (ISOPENTENYL-DIPHOSPHATE:TRNA ISOPENTENYLTRANSFERASE)
DE (IPPTASE) (IPPT).
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OX Pseudomonas.
OX NCBI_TaxId=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M.
RA Olekhnovich I.N., Gussin G.N.;
RT "Attenuation of the pseudomonas putida trpE and trpDC genes."
RL Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE FIRST STEP IN THE BIOSYNTHESIS OF
CC 2-METHYLTHIO-N6-(DELTA(2)-ISOPENTENYL)-ADENOSINE (MSI[2]I[6]A1)
CC ADJACENT TO THE ANTICODON OF SEVERAL TRNA SPECIES (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ISOPENTENYL DIPHOSPHATE + TRNA -
CC PYROPHOSPHATE + TRNA CONTAINING 6-ISOPENTENYLADENOSINE.
CC -1- SIMILARITY: BELONGS TO THE IPP TRANSFERASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF016312; AAB69443.1; -
DR InterPro: IPR002627; IPPT.
DR Pfam: PF01715; IPPT; 1.
DR PRODOM: PD004674; IPPT; 1.
KM Transferase: Nucleotidyltransferase; tRNA processing; ATP-binding.
FT NP_BIND 12 19 ATP (POTENTIAL).

SQ SEQUENCE 322 AA; 3548 MW; DIA81A2D7B32F6A CRC64;
 Query Match 1.6%; Score 7; DB 1; Length 322;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 33 FVDEVRS 39
 Db 231 FVDEVRS 237
 RESULT 10
 SPL_BACBR STANDARD; PRT; 326 AA.
 ID SPL_BACBR
 AC P43131;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE PROTEASE INHIBITOR PRECURSOR (BBRP1).
 OS Bacillus brevis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 CC Bacillus/Staphylococcus group; Brevibacillus.
 NX NCBI_TaxID=1393;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-32; 104-112 AND 122-136.
 RC STRAIN=HPD31;
 RX MEDLINE=92304060; PubMed=1610177;
 RA Shiga Y., Hasegawa K., Tsuboi A., Yamagata H., Ueda S.;
 RT "Characterization of an extracellular protease inhibitor of Bacillus
 RT brevis HPD31 and nucleotide sequence of the corresponding gene.";
 RL Appl. Environ. Microbiol. 58:525-531(1992).
 CC -1- FUNCTION: SHOWS INHIBITORY ACTIVITY TOWARDS SERINE PROTEASES, SUCH
 CC AS TRYPSIN, CHYMOTRYPSIN, AND SUBTILISIN. MAY FORM A TRYPSIN-
 CC INHIBITOR COMPLEX IN A MOLAR RATIO OF 1:1. IT IS HEAT RESISTANT AT
 CC NEUTRAL AND ACIDIC PHs.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- PTM: MAY UNDERGO ACTIVATION AFTER SECRETION. IT IS PRODUCED.
 CC EXTRACELLULARLY IN MULTIPLE FORMS HAVING AT LEAST THREE DIFFERENT
 CC MOLECULAR WEIGHTS (BBRP1-A, -B, AND -C).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D10696; BAA01538.1; -
 KW Serine protease inhibitor; Protease inhibitor; Signal; Repeat.
 FT SIGNAL 1 24
 FT CHAIN 25 326 SERINE PROTEASE INHIBITOR.
 FT CHAIN 104 326 SERINE PROTEASE INHIBITOR-C.
 FT CHAIN 104 ? SERINE PROTEASE INHIBITOR-B (POTENTIAL).
 FT CHAIN 122 326 SERINE PROTEASE INHIBITOR-A.
 FT DOMAIN 177 304 CONTAINS TWO APPROXIMATE REPEATS.
 FT REPEAT 177 208 1.
 FT REPEAT 272 304 2.
 SO SEQUENCE 326 AA; 35100 MW; 1C0456ABFA912F77 CRC64;
 Query Match 1.6%; Score 7; DB 1; Length 326;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 30 PVAFVDE 36
 Db 119 PVAFVDE 125
 RESULT 11
 FLAD_METJA
 FLAD_METJA

ID FLAD_METJA STANDARD; PRT; 342 AA.
 AC Q36305;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PUTATIVE FLAGELLA-RELATED PROTEIN D.
 GN FLAD OR MJ0895.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
 CC Methanococcus.
 NX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Usterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kalne B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT jannaschii.";
 RL Science 273:1058-1073(1996).
 CC -1- SIMILARITY: STRONG. TO M.VOLTAE FLAD, ALSO TO FLAE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U67533; AAB98898.1; -
 DR TIGR: MJ0895; -
 KW Flagella; Complete proteome.
 SO SEQUENCE 342 AA; 39950 MW; B384DDDE1775566C CRC64;
 Query Match 1.6%; Score 7; DB 1; Length 342;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 356 PDILDY 362
 Db 255 PDILDY 261
 RESULT 12
 Y755_METJA STANDARD; PRT; 342 AA.
 ID Y755_METJA
 AC O58165;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOTHETICAL PROTEIN MJ0755 PRECURSOR.
 GN MJ0755.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
 CC Methanococcus.
 NX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Usterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kalne B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT jannaschii.";
 RL Science 273:1058-1073(1996).
 CC -1- SIMILARITY: STRONG. TO M.VOLTAE FLAD, ALSO TO FLAE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U67533; AAB98898.1; -
 DR TIGR: MJ0895; -
 KW Flagella; Complete proteome.
 SO SEQUENCE 342 AA; 39950 MW; B384DDDE1775566C CRC64;

RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotten M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Kent H.-P., Fraser C.M., Smith H.O., Moese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, *Methanococcus*
 RT *jannaschii*."; Science 273:1058-1073(1996).

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U67521; AAB98748.1; -
 DR TIGR: M10755; -
 KW Hypothetical protein; Signal; Complete proteome.
 FT SIGNAL 18
 FT CHAIN 19 342
 FT HYPOTHETICAL PROTEIN M10755.
 SQ SEQUENCE 342 AA; 38778 MW; BD25220A7EAD85C CRC64;

Query Match 1.6%; Score 7; DB 1; Length 342;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 76 YTTALEN 82
 |||||
 Db 90 YTTALEN 96

RESULT 13
 NU2M_CAPHI STANDARD; PRT; 347 AA.
 AC 036346;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 2 (EC 1.6.5.3).
 GN MTND2 OR ND2.
 OS Capra hircus (Goat).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Capra.
 OX NCBI_TaxID=9925;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Doye P., Mann W., Hecht W.;
 RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE - NAD(+) + UBIQUINOL.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 CC INNER MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 2 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X72965; CA51468.1; -
 DR InterPro: IPR003917; NADHUb_oxidatcse2.
 DR InterPro: IPR001750; Oxidored_g1.
 DR Pfam: PF00361; oxidored_g1; 1.
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
 SQ SEQUENCE 347 AA; 39136 MW; 6129CFF90980475E CRC64;

Query Match 1.6%; Score 7; DB 1; Length 347;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 LTLSTLS 14
 |||||
 Db 155 LTLSTLS 161

RESULT 14
 NU2M_HUMAN STANDARD; PRT; 347 AA.
 ID NU2M_HUMAN
 AC P03891; Q34769; Q9TGI0; Q9TGI1; Q9TGI2; Q9TGI3; Q9TGI4;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 2 (EC 1.6.5.3).
 GN MTND2 OR ND2.
 OS Homo sapiens (Human).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81173052; PubMed=7219534;
 RA Anderson S., Bankier A.T., Barrell B.G., de Bruijn M.H.L.,
 RA Coulson A.R., Drouin J., Eperon I.C., Nierlich D.P., Roe B.A.,
 RA Sanger F., Schreier P.H., Smith A.J.H., Staden R., Young I.G.;
 RT "Sequence and organization of the human mitochondrial genome."; Nature 290:457-465(1981).
 RL Nature 290:457-465(1981).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81170577; PubMed=6260957;
 RA Sanger F., Coulson A.R., Barrell B.G., Smith A.J.H., Roe B.A.;
 RT "Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing."; J. Mol. Biol. 143:161-178(1980).
 RL J. Mol. Biol. 143:161-178(1980).
 RN [3]
 RP SEQUENCE FROM N.A., AND VARIANTS.
 RX MEDLINE=9813898; PubMed=9475751;
 RA Wise C.A., Strahl M., Eastaugh S.;
 RT "Departure from neutrality at the mitochondrial NADH dehydrogenase
 RT subunit 2 gene in humans, but not in chimpanzees."; Genetics 148:409-421(1998).
 RL Genetics 148:409-421(1998).
 RN [4]
 RP SEQUENCE FROM N.A., AND VARIANTS ILE-43; LEU-325 AND THR-331.
 RC TISSUE=Placenta;
 RX MEDLINE=95152634; PubMed=7530363;
 RA Horai S., Hayasaka K., Kondo R., Tsugane K., Takahata N.;
 RT "Recent African origin of modern humans revealed by complete sequences
 RT of hominoid mitochondrial DNAs."; Proc. Natl. Acad. Sci. U.S.A. 92:532-536(1995).
 RL Proc. Natl. Acad. Sci. U.S.A. 92:532-536(1995).
 RN [5]
 RP IDENTIFICATION OF PROTEIN.
 RX MEDLINE=85188293; PubMed=3921850;
 RA Chomyn A., Mariotti P., Cleeter M.W.J., Ragan C.I., Matsuno-Yagi A.,
 RA Hattel Y., Doolittle R.F., Attardi G.;
 RT "Six unidentified reading frames of human mitochondrial DNA encode
 RT components of the respiratory-chain NADH dehydrogenase."; Nature 314:592-597(1985).
 RL Nature 314:592-597(1985).
 RN [6]
 RP VARIANT LHON ASP-150.
 RX MEDLINE=91144615; PubMed=1900003;
 RA Johns D.R., Berman J.;
 RT "Alternative, simultaneous complex I mitochondrial DNA mutations in
 RT Leber's hereditary optic neuropathy."; Biochem. Biophys. Res. Commun. 174:1324-1330(1991).
 RL [7]
 RP VARIANTS L-42; R-63; A-119; P-148; S-150; T-159 AND A-185.
 RX MEDLINE=92098084; PubMed=1757091;
 RA Marzuki S., Noer A.S., Lertit P., Thyagarajan D., Kapsa R.,
 RA Utthanaphol P., Byrne E.;

RT "Normal variants of human mitochondrial DNA and translation products:
RT the building of a reference data base.";
RL Hum. Genet. 88:139-145(1991).
RN [8]
RX VARIANT LHON SER-259.
RX MEDLINE=92120513; Pubmed-1732158;
RA Brown M.D., Voljavec A.S., Lott M.T., Torroni A., Yang C.C.,
RA Wallace D.C.;
RT "Mitochondrial DNA complex I and III mutations associated with
RT Leber's hereditary optic neuropathy.";
RL Genetics 130:163-173(1992).
RN [9]
RX VARIANT AD SER-331.
RX MEDLINE=92118019; Pubmed-1370613;
RA Lin F.-H., Lin R., Wisniewski H.M., Hwang Y.-W., Grundke-Iqbal I.,
RA Healy-Louie G., Iqbal K.;
RT "Detection of point mutations in codon 331 of mitochondrial NADH
RT dehydrogenase subunit 2 in Alzheimer's brains.";
RL Biochem. Biophys. Res. Commun. 182:238-246(1992).
RN [10]
RX VARIANT THR-57.
RX MEDLINE=9812794; Pubmed-9461455;
RA Rieder M.J., Taylor S.L., Tobe V.O., Nickerson D.A.;
RT "Automating the identification of DNA variations using quality-based
RT fluorescence re-sequencing: analysis of the human mitochondrial
RT genome.";
RL Nucleic Acids Res. 26:967-973(1998).
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE.
CC -1- DISEASE: DEFECTS IN MTND2 ARE ONE OF THE CAUSES OF LEBER'S
CC HEREDITARY OPTIC NEUROPATHY (LHON), A MATERNALLY INHERITED DISEASE
CC RESULTING IN ACUTE BILATERAL BLINDNESS DUE TO RETINAL DEGENERATION
CC PREDOMINANTLY IN YOUNG MEN. CARDIAC CONDUCTION DEFECTS AND
CC NEUROLOGICAL DEFECTS HAVE ALSO BEEN DESCRIBED. RESULTING IN OPTIC
CC NERVE DEGENERATION AND CARDIAC DYSRHYTHMIA.
CC -1- DISEASE: DEFECTS IN MTND2 COULD BE ASSOCIATED WITH ALZHEIMER'S
CC DISEASE (AD).
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: J01415; AAB58944.1; -;
DR EMBL: V00662; CAA24027.1; -;
DR EMBL: M10546; AAA65502.1; ALT_INIT.
DR EMBL: D38112; BAA07291.1; -;
DR EMBL: AF014882; AAC25441.1; -;
DR EMBL: AF014884; AAC25443.1; -;
DR EMBL: AF014885; AAC25444.1; -;
DR EMBL: AF014887; AAC25446.1; -;
DR EMBL: AF014889; AAC25448.1; -;
DR EMBL: AF014890; AAC25449.1; -;
DR EMBL: AF014891; AAC25450.1; -;
DR EMBL: AF014892; AAC25451.1; -;
DR EMBL: AF014895; AAC25454.1; -;
DR EMBL: AF014896; AAC25455.2; -;
DR EMBL: AF014897; AAC25456.1; -;
DR EMBL: AF014898; AAC25457.1; -;
DR EMBL: AF014899; AAC25458.2; -;
DR EMBL: AF014900; AAC25459.1; -;
DR EMBL: AF014901; AAC25460.1; -;
DR PIR: A00414; DNHUN2.
DR MIM: 502500; -;
DR MIM: 516001; -;
DR MIM: 535000; -;
DR InterPro: IPR003917; NADHub_oxidctase2.
DR InterPro: IPR001750; Oxidored_g1.

DR Pfam: PF00361; oxidored_g1.1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane;
KW Disease mutation; Leber's hereditary optic neuropathy;
KW Alzheimer's disease; Polymorphism.
FT VARIANT 42 42
FT P -> L.
FT /FTid=VAR_008590.
FT V -> I.
FT /FTid=VAR_011348.
FT I -> T.
FT /FTid=VAR_008591.
FT Q -> R.
FT /FTid=VAR_008592.
FT I -> V.
FT /FTid=VAR_011349.
FT N -> S.
FT /FTid=VAR_011350.
FT T -> A.
FT /FTid=VAR_008593.
FT S -> P.
FT /FTid=VAR_008594.
FT N -> D (IN LHON; SECONDARY MUTATION; DOES
FT NOT SEEM TO DIRECTLY CAUSE THE DISEASE).
FT /FTid=VAR_004755.
FT N -> S.
FT /FTid=VAR_008595.
FT I -> T.
FT /FTid=VAR_008596.
FT T -> A.
FT /FTid=VAR_008597.
FT L -> M.
FT /FTid=VAR_011351.
FT G -> S (IN LHON; RARE PRIMARY MUTATION).
FT /FTid=VAR_004756.
FT A -> T.
FT /FTid=VAR_011352.
FT A -> V.
FT /FTid=VAR_011353.
FT I -> T.
FT /FTid=VAR_011354.
FT F -> L.
FT /FTid=VAR_011355.
FT A -> S (IN AD).
FT /FTid=VAR_004758.
FT A -> T.
FT /FTid=VAR_004757.
FT T -> A.
FT /FTid=VAR_011356.
FT C06FD982317C3F2D CRC64;
SQ SEQUENCE 347 AA; 38961 MM; C06FD982317C3F2D CRC64;

Query Match 1.6%; Score 7; DB 1; Length 347;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 LTLSTLS 14
Db 155 LTLSTLS 161

RESULT 15
NU2M_PANTR
ID NU2M_PANTR STANDARD; PRT; 347 AA.
AC 021798;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 2 (EC 1.6.5.3).
GN MTND2 OR ND2.
OS Pan troglodytes (Chimpanzee).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
OX NCBI_Taxid=9598;

```
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE:98133898; PubMed:9475751;
RA Wise C.A., Strahl M., Eastrel S.;
RT "Departure from neutrality at the mitochondrial NADH dehydrogenase
RL subunit 2 gene in humans, but not in chimpanzees.";
RL Genetics 148:409-421(1998).
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL: AF014908; AAC25467.1; -
DR EMBL: AF014909; AAC25468.1; -
DR EMBL: AF014910; AAC25469.1; -
DR EMBL: AF014911; AAC25470.1; -
DR EMBL: AF014912; AAC25471.1; -
DR EMBL: AF014913; AAC25472.1; -
DR EMBL: AF014914; AAC25473.1; -
DR EMBL: AF014915; AAC25474.1; -
DR EMBL: AF014916; AAC25475.1; -
DR EMBL: AF014917; AAC25476.1; -
DR EMBL: AF014918; AAC25477.1; -
DR EMBL: AF014919; AAC25478.1; -
DR EMBL: AF014920; AAC25479.1; -
DR EMBL: AF014921; AAC25480.1; -
DR InterPro: IPR003917; NADHUb_oxidrcse2.
DR InterPro: IPR001750; Oxidored_q1.
DR Pfam: PF00361; Oxidored_q1; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
FT VARIANT 94 94 P -> S (IN STRAIN A-292).
SQ SEQUENCE 347 AA; 39020 MW; 2E8269D105810D4E CRC64;
```

```
Query Match 1.6%; Score 7; DB 1; Length 347;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 8 LTLSTLS 14
| | | | | | |
Db 155 LTLSTLS 161
```

Search completed: May 9, 2002, 13:05:53
Job time: 608 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 13:05:18 : Search time 77.39 Seconds
(without alignments)
835,410 Million cell updates/sec

Title: US-09-787-083-6

Sequence score: 442
1 MKVSLSTLTSLTSCPAIA.....YHNEATSGVGIMLDMWGL 442

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 473505 seqs, 146272329 residues

Word size: 0

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

- 1: SPREMBL_17:*
- 2: sp_Archaea:*
- 3: sp_Bacteria:*
- 4: sp_Fungi:*
- 5: sp_Human:*
- 6: sp_Invertebrate:*
- 7: sp_Mammal:*
- 8: sp_Mhc:*
- 9: sp_Organelle:*
- 10: sp_Phage:*
- 11: sp_Plant:*
- 12: sp_Rodent:*
- 13: sp_Virus:*
- 14: sp_Vertebrate:*
- 15: sp_Unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	2.0	382	2 09K007	09K007 neisseria m
2	9	2.0	409	2 09J721	09J721 neisseria m
3	9	2.0	862	2 09CU62	09CU62 pasteurella
4	9	2.0	1432	10 09AX85	09AX85 oryza sativ
5	8	1.8	299	10 09M268	09M268 arabidopsi
6	8	1.8	306	2 09CU22	09CU22 pasteurella
7	7	1.6	55	12 067608	067608 tomato gold
8	7	1.6	72	12 0912S5	0912S5 potato viru
9	7	1.6	72	12 0912S4	0912S4 potato viru
10	7	1.6	95	12 0912S3	0912S3 potato viru
11	7	1.6	96	12 0912S8	0912S8 potato viru
12	7	1.6	100	3 012148	012148 saccharomy
13	7	1.6	101	12 0912S6	0912S6 potato viru
14	7	1.6	116	2 09C1E9	09C1E9 lactococcus
15	7	1.6	122	2 09R016	09R016 delinococcus
16	7	1.6	124	12 0912S7	0912S7 potato viru
17	7	1.6	131	6 09GMD4	09GMD4 macaca fusc
18	7	1.6	138	12 0912C7	0912C7 potato viru
19	7	1.6	147	5 036035	036035 hexamita in

20	7	1.6	158	2 0929C1	0929C1 chlamydia p
21	7	1.6	169	2 09L2A7	09L2A7 streptomyce
22	7	1.6	189	13 09DF72	09DF72 notochemia
23	7	1.6	198	4 09BUX9	09BUX9 homo sapien
24	7	1.6	212	13 091819	091819 ictalurus p
25	7	1.6	214	5 09VTR5	09VTR5 drosophila
26	7	1.6	216	2 09X6H6	09X6H6 streptococc
27	7	1.6	218	2 09R1F4	09R1F4 streptococc
28	7	1.6	219	5 017594	017594 caenorhabdl
29	7	1.6	220	2 09XCK1	09XCK1 streptococc
30	7	1.6	220	5 036036	036036 spironucleu
31	7	1.6	221	12 065002	065002 alfalfa mos
32	7	1.6	227	5 09BWP4	09BWP4 glossina mo
33	7	1.6	233	6 09WZB0	09WZB0 ovis aries
34	7	1.6	236	13 09DF58	09DF58 notochemia
35	7	1.6	264	12 085260	085260 potato viru
36	7	1.6	265	9 038155	038155 bacterioph
37	7	1.6	267	9 099363	099363 bacterioph
38	7	1.6	267	12 085261	085261 potato viru
39	7	1.6	267	12 098628	098628 potato viru
40	7	1.6	267	12 086842	086842 potato viru
41	7	1.6	267	12 099026	099026 potato viru
42	7	1.6	267	12 005476	005476 potato viru
43	7	1.6	267	12 009711	009711 potato viru
44	7	1.6	267	12 09WH17	09WH17 potato viru
45	7	1.6	267	12 09QNM9	09QNM9 potato viru

ALIGNMENTS

RESULT 1
ID 09K007 PRELIMINARY; PRT: 382 AA.
AC 09K007;
DT 01-OCR-2000 (TREMBLrel. 15, Created)
DT 01-OCR-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PHOSPHOLIPASE A1, PUTATIVE.
GN NMB0464.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; Pubmed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E., Eisen J.A., Ketchum K.A., Hood B.W., Peden J.F., Dodson R.J., Nelson W.C., Gwinn M.L., Deboy R., Peterson J.D., Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A., Mason T., Clecko A., Parksey D.S., Blair E., Ciftone H., Clark E.B., Cotton M.D., Uitterlind T.R., Khouri H., Qin H., Yamathavan J., Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RA "Complete genome sequence of Neisseria meningitidis serogroup B strain RT MC58.";
RL Science 287:1809-1815(2000).
DR EMBL: AE002403; AAF40901.1; -;
DR TIGR: NMB0464; -;
DR InterPro: IPR003187; PLAI.
DR Pfam: PF02253; PLAI; 1.
KW Complete proteome.
SQ SEQUENCE 382 AA; 42714 MW; B468A802F062E836 CRC64;

Query Match 2.0%; Score 9; DB 2; Length 382;
Best Local Similarly 100.0%; Pred.No. 1.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 279 PEIRLOPV 287
DB 221 PEIRLOPV 229

```

RESULT 2
ID 09JUT21 PRELIMINARY: PRT: 409 AA.
AC 09JUT21:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PUTATIVE PHOPHOLIPASE.
GN MMA2021.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltham T., Hamlin N., Holroyd S.,
RA Jagers K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RA "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491."
RL Nature 404:502-506 (2000).
DR EMBL; AL162757; CAB85240.1;
DR InterPro: IPR003187; Pfam.
DR Pfam: PF02253; Pfam; 1.
KW Complete proteome.
SQ SEQUENCE 409 AA; 45862 MW; CD6585B064D01A41 CRC64;

Query Match
Best Local Similarity 2.0%; Score 9; DB 2; Length 409;
Matches 9; Conservative 100.0%; Pred. No. 1.4; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIFLTQPV 287
Db 248 PEIFLTQPV 256

RESULT 3
ID 09CL62 PRELIMINARY: PRT: 862 AA.
AC 09CL62:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CVG81.
GN CVG81 OR PM1380.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida PM70."
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465 (2001).
CC -1- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
CC REGULATORY COMPONENTS OF SENSORY TRANSDUCTION SYSTEMS.
CC -1- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
CC KINASES.
DR EMBL; AF006176; AA033464.1;
DR InterPro: IPR000410; Bctrl_sensory.
DR InterPro: IPR000658; DUF5.
DR InterPro: IPR003660; HAMF.
DR InterPro: IPR003594; HATPase_C.
DR InterPro: IPR003661; His_kinA.

```

```

DR InterPro: IPR001789; Response_reg.
DR Pfam: PF00672; DUF5; 1.
DR Pfam: PF00072; response_reg; 1.
DR Pfam: PF00512; signal; 1.
DR PRINTS: PR00344; BCTRLSENSOR.
DR SMART: SM00304; HAMF; 1.
DR SMART: SM00387; HATPase_C; 1.
DR SMART: SM00388; HSKA; 1.
DR SMART: SM00448; REC; 1.
KW Complete proteome; Kinase; Phosphorylation; Sensory transduction;
KW Transferase.
SQ SEQUENCE 862 AA; 97846 MW; AB681A3B6326C0 CRC64;

Query Match
Best Local Similarity 2.0%; Score 9; DB 2; Length 862;
Matches 9; Conservative 100.0%; Pred. No. 2.7; Mismatches 0; Indels 0; Gaps 0;

QY 5 LSTLTSL 13
Db 293 LSTLTSL 301

RESULT 4
ID 09AX85 PRELIMINARY: PRT: 1432 AA.
AC 09AX85:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PUTATIVE ABC TRANSPORTER PROTEIN.
GN P0410E03.7.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartiaceae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, PAC
RL clone: P0410E03."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002844; BAB21276.1;
SQ SEQUENCE 1432 AA; 160234 MW; 04E8CF2CA08D993 CRC64;

Query Match
Best Local Similarity 2.0%; Score 9; DB 10; Length 1432;
Matches 9; Conservative 100.0%; Pred. No. 4.2; Mismatches 0; Indels 0; Gaps 0;

QY 170 FTPLSLFD 178
Db 830 FTPLSLFD 838

RESULT 5
ID 09M268 PRELIMINARY: PRT: 299 AA.
AC 09M268:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE RAV-LIKE PROTEIN.
GN F21F14.140.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN (1)
RP SEQUENCE FROM N.A.

```

RA Choiene N., Robert C., Brotlier P., Wincker P., Catolico L.,
RA Artiguenave F., Saurin W., Weissenbach J., Kewes H.W., Lemcke K.,
RA Mayer K.F.X., Quelier F., Salanoubat M.,
RA Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL138642; CAB71904.1; -.
DR InterPro: IPR003340; B3.
DR Pfam: PF02362; B3; 1.
SQ SEQUENCE 299 AA; 34270 MW; 2AB841F179DD174B CRC64;

Query Match 1.8%; Score 8; DB 10; Length 299;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 174 SLSPDLDR 181
DB 292 SLSPDLDR 299

RESULT 6
Q9CL22 PRELIMINARY; PRT; 306 AA.
AC Q9CL22;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE HYPOTHETICAL PROTEIN PM1426.
GN PM1426.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=PM70;
RC MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL: AE006179; AK03510.1; -.
DR InterPro: IPR003187; PLA1.
DR Pfam: PF02253; PLA1; 1.
KM Hypothetical protein: Complete proteome.
SQ SEQUENCE 306 AA; 35580 MW; EAF3DE8C1C22B26E CRC64;

Query Match 1.8%; Score 8; DB 2; Length 306;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 314 KLSRSNR 321
DB 184 KLSRSNR 191

RESULT 7
Q67608 PRELIMINARY; PRT; 55 AA.
AC Q67608;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE SUBGENOMIC DNA DERIVED FROM DNA B CCCDS - COVALENTLY CLOSED CIRCULAR
DE DOUBLE-STRANDED MOLECULE.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10831;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=87040767; PubMed=3022243;
RA McDowell S.W., Coutts R.H.A., Buck K.W.;
RT "Molecular characterisation of subgenomic single-stranded and double-
RT stranded DNA forms isolated from plants infected with tomato golden
RT mosaic virus.";
RL Nucleic Acids Res. 14:7967-7984(1986).
DR EMBL: X04485; CAA28171.1; -.
DR InterPro: IPR000211; Gemin1_B1.
DR Pfam: PF00845; Gemin1_B1; 1.
SQ SEQUENCE 55 AA; 6281 MW; 41963570D739EC96 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 55;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 312 SAKLSRS 318
DB 43 SAKLSRS 49

RESULT 8
Q91ZS5 PRELIMINARY; PRT; 72 AA.
AC Q91ZS5;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y strain NTN (PVY(NTN)).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyviruses.
OX NCBI_TaxID=122280;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=TUBER NECROSING (NTN);
RA Czerovska N., Filigiarova M., Moravec T., Petrzik K.;
RT "Differences in nucleotide and amino acid sequences of N-terminal
RT parts of coat proteins among isolates of potato virus Y-NTN strain.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF228633; AAF67874.1; -.
FT NON_TER 1 1
FT NON_TER 72 72
SQ SEQUENCE 72 AA; 7721 MW; 2C5287E15624CBDD CRC64;

Query Match 1.6%; Score 7; DB 12; Length 72;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 156 TKKDAKQ 162
DB 11 TKKDAKQ 17

RESULT 9
Q91ZS4 PRELIMINARY; PRT; 72 AA.
AC Q91ZS4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y strain NTN (PVY(NTN)).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyviruses.
OX NCBI_TaxID=122280;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=TUBER NECROSING (NTN);
RA Czerovska N., Filigiarova M., Moravec T., Petrzik K.;

RT "Differences in nucleotide and amino acid sequences of N-terminal
RT parts of coat proteins among isolates of potato virus Y-NTN strain."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF228634; AAF67875.1; -
FT NON_TER 1 1
FT SEQUENCE 72 AA; 7692 MW; 4F129B870523B482 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 72;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKO 162
DB 11 TKKDAKO 17

RESULT 10

ID 091ZS3 PRELIMINARY; PRT; 95 AA.
AC 091ZS3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y strain NTN (PVY(NTN)).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Polyviruses.
OX NCBI_TaxID=122280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUBER NECROSING (NTN);
RA Czerovska N., Filigayova M., Moravec T., Petrzik K.;
RT "Differences in nucleotide and amino acid sequences of N-terminal
RT parts of coat proteins among isolates of potato virus Y-NTN strain."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF228635; AAF67876.1; -
FT NON_TER 1 1
FT SEQUENCE 95 AA; 10446 MW; 28C06891777C9928 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 95;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKO 162
DB 11 TKKDAKO 17

RESULT 11

ID 091ZS8 PRELIMINARY; PRT; 96 AA.
AC 091ZS8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y (strain N) (PVY).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Polyviruses.
OX NCBI_TaxID=12219;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NECROSING (N);
RA Czerovska N., Filigayova M., Moravec T., Petrzik K.;
RT "Differences in nucleotide and amino acid sequences of N-terminal
RT parts of coat proteins among isolates of potato virus Y-NTN strain."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF228630; AAF67871.1; -
FT NON_TER 1 1
FT SEQUENCE 96 AA; 10559 MW; 89C3A5675B9784CA CRC64;

Query Match 1.6%; Score 7; DB 12; Length 96;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKO 162
DB 11 TKKDAKO 17

RESULT 12

ID 012148 PRELIMINARY; PRT; 100 AA.
AC 012148;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE ORF YDL163W.
GN YDL163W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Pohl T.M.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MIPS;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ALPHAS288C;
RA Pohl T.M.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; 274212; CAA98738.1; -
DR EMBL; 267750; CAA91583.1; -
DR SGD; S0002322; YDL163W.
KW Hypothetical protein.
SQ SEQUENCE 100 AA; 10751 MW; 952FB92C852AA710 CRC64;

Query Match 1.6%; Score 7; DB 3; Length 100;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 LFLSILS 14
DB 26 LFLSILS 32

RESULT 13

ID 091ZS6 PRELIMINARY; PRT; 101 AA.
AC 091ZS6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y strain NTN (PVY(NTN)).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Polyviruses.
OX NCBI_TaxID=122280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUBER NECROSING (NTN);

RA Carovska N., Filigajrova M., Moravec T., Petrzik K.;
RT "Differences in nucleotide and amino acid sequences of N-terminal
RT parts of coat proteins among isolates of potato virus Y-NTN strain."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF228632; AAF67873.1; -
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 101 AA; 11112 MW; B55BP1C3D5FE9BD CRC64;

Query Match 1.6%; Score 7; DB 12; Length 101;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKQ 162
Db 11 TKKDAKQ 17

RESULT 14
ID Q9CIE9 PRELIMINARY; PRT; 116 AA.
AC Q9CIE9;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE CELLOBIOSE-SPECIFIC PTS SYSTEM IIA COMPONENT (EC 2.7.1.69).
GN PTCA.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria: Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus;
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RA Bolotin A., Wincker P., Manger S., Jallion O., Malarne K.,
RA Weissensbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis."
RL Genome Res. 0:0-0(2001).
DR EMBL: AE006278; AK04512.1; -
DR InterPro: IPR003188; PTS_IIA.
DR Pfam: PF02255; PTS_IIA; 1.
KW Complete proteome.
SQ SEQUENCE 116 AA; 12814 MW; 995F034417DA3DA0 CRC64;

Query Match 1.6%; Score 7; DB 2; Length 116;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 AIOQAKA 26
Db 34 AIOQAKA 40

RESULT 15
Q9RU16
ID Q9RU16 PRELIMINARY; PRT; 122 AA.
AC Q9RU16;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE HYPOTHETICAL 13.7 KDA PROTEIN.
GN DR1580.
OS Deinococcus radiodurans.
OC Bacteria: Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RI;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,

RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans RI."
RL Science 286:1571-1577(1999).
DR EMBL: AE002002; AAF1150.1; -
DR TIGR: DR1580; -
DR InterPro: IPR000325; Glyoxalase_1.
DR Pfam: PF00903; Glyoxalase; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 122 AA; 13683 MW; 7BA836DCAD786E51 CRC64;

Query Match 1.6%; Score 7; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 EHEPELY 76
Db 115 EHEPELY 121

Search completed: May 9, 2002, 13:05:18
Job time: 613 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:54:59 ; Search time 72.68 Seconds
(without alignments)
450.473 Million cell updates/sec

Title: US-09-787-083-8

Perfect score: 442
Sequence: 1 MKVSLSTLSTLPCFALA.....YHNEATSPGVGLMDNMGL 442

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 522463 seqs, 74073290 residues

Word size : 0

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

A_Geneseq_1101.*
1: /SIDSR/gcgdata/geneseq/AA1980.DAT.*
2: /SIDSR/gcgdata/geneseq/AA1981.DAT.*
3: /SIDSR/gcgdata/geneseq/AA1982.DAT.*
4: /SIDSR/gcgdata/geneseq/AA1983.DAT.*
5: /SIDSR/gcgdata/geneseq/AA1984.DAT.*
6: /SIDSR/gcgdata/geneseq/AA1985.DAT.*
7: /SIDSR/gcgdata/geneseq/AA1986.DAT.*
8: /SIDSR/gcgdata/geneseq/AA1987.DAT.*
9: /SIDSR/gcgdata/geneseq/AA1988.DAT.*
10: /SIDSR/gcgdata/geneseq/AA1989.DAT.*
11: /SIDSR/gcgdata/geneseq/AA1990.DAT.*
12: /SIDSR/gcgdata/geneseq/AA1991.DAT.*
13: /SIDSR/gcgdata/geneseq/AA1992.DAT.*
14: /SIDSR/gcgdata/geneseq/AA1993.DAT.*
15: /SIDSR/gcgdata/geneseq/AA1994.DAT.*
16: /SIDSR/gcgdata/geneseq/AA1995.DAT.*
17: /SIDSR/gcgdata/geneseq/AA1996.DAT.*
18: /SIDSR/gcgdata/geneseq/AA1997.DAT.*
19: /SIDSR/gcgdata/geneseq/AA1998.DAT.*
20: /SIDSR/gcgdata/geneseq/AA1999.DAT.*
21: /SIDSR/gcgdata/geneseq/AA2000.DAT.*
22: /SIDSR/gcgdata/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	442	100.0	442	21	AAV85271
2	327	74.0	442	21	BASB034 amino acid
3	316	71.5	442	21	AAV85268
4	289	65.4	442	21	AAV85270
5	9	2.0	370	21	AAV85269
6	9	2.0	370	21	AAV75156
7	9	2.0	370	21	AAV75157
8	9	2.0	370	21	AAV75158
9	9	2.0	374	21	AAV70629
10	8	1.8	278	21	AAV70628
11	8	1.8	283	21	AAV1137
					Arabidopsis thalia

12	8	1.8	299	21	AAV1135	Arabidopsis thalia
13	7	1.6	16	15	AAV52127	Mouse light chain
14	7	1.6	50	21	AAV79463	Rat alpha tubulin
15	7	1.6	64	21	AAV06322	Arabidopsis thalia
16	7	1.6	82	21	AAV01488	Human secreted pro
17	7	1.6	106	21	AAV07608	Arabidopsis thalia
18	7	1.6	106	21	AAV1952	Arabidopsis thalia
19	7	1.6	108	21	AAV37811	Arabidopsis thalia
20	7	1.6	116	21	AAV87313	Human signal pepti
21	7	1.6	136	21	AAV37810	Arabidopsis thalia
22	7	1.6	158	21	AAV79176	Haematopoietic ste
23	7	1.6	169	20	AAV34661	Chlamydia pneumoni
24	7	1.6	181	21	AAV37809	Arabidopsis thalia
25	7	1.6	186	21	AAV07607	Arabidopsis thalia
26	7	1.6	186	21	AAV51951	Arabidopsis thalia
27	7	1.6	188	21	AAV31950	Arabidopsis thalia
28	7	1.6	190	21	AAV07606	Arabidopsis thalia
29	7	1.6	268	13	AAV21542	Envelope protein o
30	7	1.6	302	13	AAV24254	Protease inhibitor
31	7	1.6	316	21	AAV39500	Arabidopsis thalia
32	7	1.6	324	21	AAV39489	Arabidopsis thalia
33	7	1.6	326	13	AAV24356	Protease inhibitor
34	7	1.6	365	21	AAV39498	Arabidopsis thalia
35	7	1.6	381	20	AAV60241	Human endometrium
36	7	1.6	448	21	AAV91960	Human cytoskeleton
37	7	1.6	523	21	AAV44285	Arabidopsis thalia
38	7	1.6	552	16	AAV67007	Firefly luciferase
39	7	1.6	666	20	AAV21871	Amino acid sequenc
40	7	1.6	683	20	AAV21869	Amino acid sequenc
41	7	1.6	684	20	AAV21873	Amino acid sequenc
42	7	1.6	715	20	AAV21875	Amino acid sequenc
43	7	1.6	715	20	AAV21860	Amino acid sequenc
44	7	1.6	783	20	AAV70586	Human SLAP-130. H
45	7	1.6	900	20	AAV96264	Human axlin. Homo

ALIGNMENTS

RESULT 1	
AAV85271	standard; Protein: 442 AA.
XX	
AC	AAV85271:
XX	
DT	29-JUN-2000 (first entry)
XX	
DE	BASB034 amino acid sequence #4.
XX	
KW	Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KW	vaccine; bacteriostatic; treatment; otitis media; pneumonia;
KW	sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW	hearing loss; antibacterial drug.
XX	
OS	Moraxella catarrhalis.
XX	
PN	W0200015802-A1.
XX	
PD	23-MAR-2000.
XX	
PF	14-SEP-1999; 99MO-EP06781.
XX	
PR	14-SEP-1998; 98GB-0020002.
XX	
PA	(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX	
PI	Ruelle J;
XX	
DR	WPI; 2000-271440/23.
XX	
PT	N-PSDB; AAA10703.
XX	
PT	Novel BASB034 polynucleotides and polypeptides from Moraxella
	catarrhalis used to prepare vaccines against bacterial infections

XX Claim 3; Page 69; 106pp; English
PS

This sequence represents a *Moraxella catarrhalis* BASB034 polypeptide from strain Mc2969. The invention relates to BASB034 polypeptides from *M. catarrhalis* strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034 polynucleotides and polypeptides may be employed as research reagents and material for the discovery of treatments and diagnostics for diseases, particularly human diseases. They are particularly used to diagnose and treat *M. catarrhalis* infections. They can be used for diagnosis of disease, staging of disease, or determining response of an infectious organism to drugs. The polynucleotides may be used as a source for hybridization probes, and for screening of genetic mutations, serotype, organism or strain identification, identification of mutations in BASB034 sequences, and as components of arrays which are useful for diagnostic and prognostic purposes. The polypeptides can be used to produce antibodies. The polypeptides can also be used in vaccine formulations, and to identify agonists and antagonists. The polypeptides, antibodies, agonists and antagonists (which are bacteriostatic) are used for the treatment and prevention of diseases such as otitis media in infants and children, pneumonia in elderly, sinusitis, nosocomial infections and invasive diseases, and chronic otitis media with hearing loss. The polypeptides, agonists and antagonists are also used for screening of antibacterial drugs. The BASB034 products of the invention can be used screen for new antibacterial compounds that may target resistant bacteria.

SQ Sequence 442 AA;

Query Match	100.0%	Score 442;	DB 21;	Length 442;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 442; Conservative	0;	Mismatches	0;	Indels 0;
				Gaps 0

QY	1	mkvslstltlslrpfalalaoaqaovnpnrafvdevrseuadgondehrldvqsatqsa	60
Db	1	mkvslstltlslrpfalalaoaqaovnpnrafvdevrseuadgondehrldvqsatqsa	60
QY	61	stdtjnpnlrdehebrpelyttaleenkmtmlncsalndinrlacudrlyvgefpavlyktkrsi	120
Db	61	stdtjnpnlrdehebrpelyttaleenkmtmlncsalndinrlacudrlyvgefpavlyktkrsi	120
QY	121	rldefitwotitgkrfouvyoetttdrflmgnekgmltkkdaokoleayaaakofpyslsfdd	180
Db	121	rldefitwotitgkrfouvyoetttdrflmgnekgmltkkdaokoleayaaakofpyslsfdd	180
QY	181	rnnpflmsrphnnpuyvlrpfmhnkprnspmtsehnkqotpnrfearbelkrfouvyvuka	240
Db	181	rnnpflmsrphnnpuyvlrpfmhnkprnspmtsehnkqotpnrfearbelkrfouvyvuka	240
QY	241	aedlwtgtsdmlmfgytqooshnqolfnngknsrpfryhndvorpelfiltoapyulsdpmvdkvymt	300
Db	241	aedlwtgtsdmlmfgytqooshnqolfnngknsrpfryhndvorpelfiltoapyulsdpmvdkvymt	300
QY	301	gmgnvnhnsgesaklsrskwnnraylmacgmeknltvmrtrimgrrfkesssgoipddndpild	360
Db	301	gmgnvnhnsgesaklsrskwnnraylmacgmeknltvmrtrimgrrfkesssgoipddndpild	360
QY	361	ygygdvayfvlqoleknsniscgtvrynprpsgaldolvyvrlgkigisgyfoifogxosl	420
Db	361	ygygdvayfvlqoleknsniscgtvrynprpsgaldolvyvrlgkigisgyfoifogxosl	420
QY	421	ldynheatsfvgvglmndmgl	442
Db	421	ldynheatsfvgvglmndmgl	442

RESULT	2
AAV85268	
XX	standard; Protein; 442 AA
AC	
XX	AAV85268;
XX	

DT	29-JUN-2000	(first entry)
XX		
DE	BASB034	amino acid sequence #1.
XX		
KW	Moraxella catarrhalis infection;	BASB034; diagnosis; straining;
KW	vaccine; bacteriostatic; treatment;	otitis media; pneumonia;
KW	sinusitis; nosocomial infection;	invasive disease; chronic otitis media;
KW	hearing loss; antibacterial drug.	
XX		
OS	Moraxella catarrhalis.	
XX		
PN	WO200015802-A1.	
XX		
PD	23-MAR-2000.	
XX		
PF	14-SEP-1999;	99WO-EP06781.
XX		
PR	14-SEP-1998;	98GB-0020002.
XX		
PA	(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.	
XX		
PI	Ruelle J;	
XX		
DR	WPI: 2000-271440/23.	
XX		
PT	N-PSDB; AAA10700.	
XX		
PT	Novel BASB034 polynucleotides and polypeptides from Moraxella	
PT	catarrhalis used to prepare vaccines against bacterial infections	-
XX		
PS	Claim 3; Fig 2; 106pp; English.	

Claim 3; Fig 2; 106pp; English.

This sequence represents a Moraxella catarrhalis BASB034 polypeptide from strain MC2931 (ATCC 43617). The invention relates to BASB034 polypeptides from M. catarrhalis strains MC2931, MC2908, MC2913 and MC2965. The BASB034 polynucleotides and polypeptides may be employed as research reagents and material for the discovery of treatments and diagnostics for diseases, particularly human diseases. They are particularly used to diagnose and treat M. catarrhalis infections. They can be used for diagnosis of disease, staging of disease, or determining response of an infectious organism to drugs. The polynucleotides may be used as a source for hybridization probes, and for screening of genetic mutations, serotype, organism or strain identification, identification of mutations in BASB034 sequences, and as components of arrays which are useful for diagnostic and prognostic purposes. The polypeptides can be used to produce antibodies. The polypeptides can also be used in vaccine formulations, and to identify agonists and antagonists. The polypeptides, antibodies, agonists and antagonists (which are bacteriostatic) are used for the treatment and prevention of diseases such as otitis media in infants and children, pneumonia in elderlies, sinusitis, nosocomial infections and invasive diseases, and chronic otitis media with hearing loss. The polypeptides, agonists and antagonists are also used for screening of antibacterial drugs. The BASB034 products of the invention can be used screen for new antibacterial compounds that may target resistant bacteria.

SQ Sequence 442 AA;

Query Match	74.08;	Score 327;	DB 21;	Length 442;
Best Local Similarity	99.88;	Pred. No. 0;		
Matches 427; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	15	CAIATIAIOAOAVPRPAFVEVRESENDLGGDNELPIDVGNATOSASRDTNPALDEHPE	74
Db	15	callaiaqgaqavpvpvafvevrsenldggdnelpidvgsa tgsaatcdnpldehpe	74
QY	75	LYTTALLENKTMALINCSALNODIMRLACVDTLVHGTEPAVITKRSIKRLDETIVOTIGKP	134
Db	75	lyttalenktmlincsalngdmrlacvdtlvhgetpavikrslrldetlwlilgkp	134
QY	135	QVVOETDPIFLIMENKGMILTKDKAKOLEVAAKFTPLSPFDLRNNTPLWSSRPHP	194
Db	135	qvlygetcdpflimnegenmltkkakaqlevaaqfcp lslstfldrnntplwssrhp	194


```

AC  AAY75157;
XX
XX  21-MAR-2000 (first entry)
XX
XX  Neisseria meningitidis ORF 582 protein sequence SEQ ID NO:1788.
DE
XX  Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
XX  antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
XX  antibacterial; gene therapy.
XX
XX  Neisseria meningitidis.
OS
XX
XX  WO957280-A2.
XX
XX  11-NOV-1999.
XX
XX  30-APR-1999; 99WO-US09346.
XX
XX  01-MAY-1998; 98US-0083758.
XX  31-JUL-1998; 98US-0094869.
XX  02-SEP-1998; 98US-0098994.
XX  02-SEP-1998; 98US-0099062.
XX  09-OCT-1998; 98US-0103749.
XX  09-OCT-1998; 98US-0103794.
XX  09-OCT-1998; 98US-0103796.
XX  25-FEB-1999; 99US-0121528.
XX
XX  (CHIR ) CHIRON CORP.
XX  (GENO-) INST GENOMIC RES.
XX
XX  Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M;
XX  Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
XX  Tettelin H, Venter JC;
XX
XX  WPI: 2000-062150/05.
XX  N-PSDB; AA253919.
XX
XX  Novel Neisserial polypeptides predicted to be useful antigens for
XX  vaccines and diagnostics -
XX
XX  Claim 2; Page 904; 1453pp; English.
XX
XX  AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941
XX  represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
XX  and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
XX  PCR primers used in the exemplification of the present invention. The
XX  polypeptides, the polynucleotides, antibodies and compositions of
XX  the invention can be used as vaccines, as diagnostic reagents, and as
XX  immunogenic compositions. The polypeptides can be used in the
XX  manufacture of medicaments for treating or preventing infection due to
XX  Neisserial bacteria (e.g. meningitis and septicemia), to detect the
XX  presence of Neisseria bacteria, or to raise antibodies. They may also
XX  be used to screen for agonists or antagonists, which may themselves
XX  have use as antibacterial agents. The polynucleotides of the invention
XX  may also be used in gene therapy protocols.
XX
XX
XX  Sequence 370 AA;
XX
XX  Query Match 2.0%; Score 9; DB 21; Length 370;
XX  Best Local Similarity 100.0%; Pred. No. 2.1;
XX  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX  QY 279 PEIFLTQPV 287
XX  |||||
XX  Db 209 pelfltqpv 217

```

```

XX
XX  21-MAR-2000 (first entry)
XX
XX  Neisseria meningitidis ORF 582 protein sequence SEQ ID NO:1790.
DE
XX  Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
XX  antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
XX  antibacterial; gene therapy.
XX
XX  Neisseria meningitidis.
OS
XX
XX  WO957280-A2.
XX
XX  11-NOV-1999.
XX
XX  30-APR-1999; 99WO-US09346.
XX
XX  01-MAY-1998; 98US-0083758.
XX  31-JUL-1998; 98US-0094869.
XX  02-SEP-1998; 98US-0098994.
XX  02-SEP-1998; 98US-0099062.
XX  09-OCT-1998; 98US-0103749.
XX  09-OCT-1998; 98US-0103794.
XX  09-OCT-1998; 98US-0103796.
XX  25-FEB-1999; 99US-0121528.
XX
XX  (CHIR ) CHIRON CORP.
XX  (GENO-) INST GENOMIC RES.
XX
XX  Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M;
XX  Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
XX  Tettelin H, Venter JC;
XX
XX  WPI: 2000-062150/05.
XX  N-PSDB; AA253920.
XX
XX  Novel Neisserial polypeptides predicted to be useful antigens for
XX  vaccines and diagnostics -
XX
XX  Claim 2; Page 905; 1453pp; English.
XX
XX  AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941
XX  represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
XX  and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
XX  PCR primers used in the exemplification of the present invention. The
XX  polypeptides, the polynucleotides, antibodies and compositions of
XX  the invention can be used as vaccines, as diagnostic reagents, and as
XX  immunogenic compositions. The polypeptides can be used in the
XX  manufacture of medicaments for treating or preventing infection due to
XX  Neisserial bacteria (e.g. meningitis and septicemia), to detect the
XX  presence of Neisseria bacteria, or to raise antibodies. They may also
XX  be used to screen for agonists or antagonists, which may themselves
XX  have use as antibacterial agents. The polynucleotides of the invention
XX  may also be used in gene therapy protocols.
XX
XX
XX  Sequence 370 AA;
XX
XX  Query Match 2.0%; Score 9; DB 21; Length 370;
XX  Best Local Similarity 100.0%; Pred. No. 2.1;
XX  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX  QY 279 PEIFLTQPV 287
XX  |||||
XX  Db 209 pelfltqpv 217

```

```

RESULT 7
AAY75158
ID AAY75158 standard; Protein: 370 AA.
XX
XX  AAY75158;
AC

```

```

RESULT 8
AAY70629
ID AAY70629 standard; Protein: 374 AA.
XX
XX  AAY70629;
AC

```

DT 18-JUL-2000 (first entry)
XX Neisseria meningitidis serogroup B strain H44/76 BASB033 protein.
DE
XX BASB033: diagnosis, prophylaxis; treatment; antibacterial; vaccine;
KM Neisseria meningitidis infection.
XX
OS Neisseria meningitidis.
XX
PN WO200015801-A1.
XX
PD 23-MAR-2000.
XX
PF 09-SEP-1999; 99WO-EP06718.
XX
PR 14-SEP-1998; 98GB-0020003.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Ruelle J;
XX
DR WPI: 2000-271439/23.
XX
DR N-PSDB; AA252134.
XX
PT Isolated BASB033 polypeptides and polynucleotides of Neisseria
PT meningitidis, useful for diagnosis, prophylaxis and treatment of N.
XX meningitidis infection -
XX
PS Claim 4; Page 59; 93pp; English.
XX
CC The present sequence is a BASB033 protein from
CC Neisseria meningitidis serogroup B strain H44/76. The protein
CC shows homology to the Klebsiella pneumoniae outer membrane
CC phospholipase A. The present sequence is useful for diagnosis,
CC prophylaxis and treatment of N. meningitidis infection. It may also be
CC used for the discovery and development of antibacterial compounds and
CC in vaccine compositions.
XX
SQ Sequence 374 AA;

Query Match 2.0%; Score 9; DB 21; Length 374;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIFLTQPV 287
| | | | | | | | | |
Db 213 peifltqpv 221

RESULT 9
AAV70628
ID AAV70628 standard; Protein; 375 AA.
XX
AC AAV70628;
XX
DT 18-JUL-2000 (first entry)
XX
DE Neisseria meningitidis serogroup B strain ATCC13090 BASB033 protein.
XX
KM BASB033: diagnosis, prophylaxis; treatment; antibacterial; vaccine;
KM Neisseria meningitidis infection.
XX
OS Neisseria meningitidis.
XX
PN WO200015801-A1.
XX
PD 23-MAR-2000.
XX
PF 09-SEP-1999; 99WO-EP06718.
XX
PR 14-SEP-1998; 98GB-0020003.
XX

PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Ruelle J;
XX
DR WPI: 2000-271439/23.
DR N-PSDB; AA252133.
XX
XX Isolated BASB033 polypeptides and polynucleotides of Neisseria
PT meningitidis, useful for diagnosis, prophylaxis and treatment of N.
PT meningitidis infection -
XX
PS Claim 4; Page 58; 93pp; English.
XX
CC The present sequence is a BASB033 protein from
CC Neisseria meningitidis serogroup B strain ATCC13090. The protein
CC shows homology to the Klebsiella pneumoniae outer membrane
CC phospholipase A. The present sequence is useful for diagnosis,
CC prophylaxis and treatment of N. meningitidis infection. It may also be
CC used for the discovery and development of antibacterial compounds and
CC in vaccine compositions.
XX
SQ Sequence 375 AA;

Query Match 2.0%; Score 9; DB 21; Length 375;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIFLTQPV 287
| | | | | | | | | |
Db 214 peifltqpv 222

RESULT 10
AAG1137
ID AAG1137 standard; Protein; 278 AA.
XX
AC AAG1137;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 9741.
XX
KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 01-APR-1999; 99US-0126785.
PR 06-APR-1999; 99US-0127462.
PR 08-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 28-APR-1999; 99US-0130510.
PR 30-APR-1999; 99US-0130691.
PR 04-MAY-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999;	99US-0132485.
PR 06-MAY-1999;	99US-0132486.
PR 06-MAY-1999;	99US-0132487.
PR 07-MAY-1999;	99US-0132863.
PR 11-MAY-1999;	99US-0134256.
PR 14-MAY-1999;	99US-0134218.
PR 14-MAY-1999;	99US-0134219.
PR 14-MAY-1999;	99US-0134221.
PR 14-MAY-1999;	99US-0134370.
PR 18-MAY-1999;	99US-0134768.
PR 19-MAY-1999;	99US-0134941.
PR 20-MAY-1999;	99US-0135124.
PR 21-MAY-1999;	99US-0135353.
PR 24-MAY-1999;	99US-0135629.
PR 25-MAY-1999;	99US-0136021.
PR 27-MAY-1999;	99US-0136392.
PR 28-MAY-1999;	99US-0136782.
PR 01-JUN-1999;	99US-0137222.
PR 03-JUN-1999;	99US-0137528.
PR 04-JUN-1999;	99US-0137502.
PR 07-JUN-1999;	99US-0137724.
PR 08-JUN-1999;	99US-0138094.
PR 10-JUN-1999;	99US-0138540.
PR 10-JUN-1999;	99US-0138847.
PR 14-JUN-1999;	99US-0139119.
PR 16-JUN-1999;	99US-0139452.
PR 16-JUN-1999;	99US-0139453.
PR 17-JUN-1999;	99US-0139492.
PR 18-JUN-1999;	99US-0139454.
PR 18-JUN-1999;	99US-0139455.
PR 18-JUN-1999;	99US-0139456.
PR 18-JUN-1999;	99US-0139457.
PR 18-JUN-1999;	99US-0139458.
PR 18-JUN-1999;	99US-0139459.
PR 18-JUN-1999;	99US-0139460.
PR 18-JUN-1999;	99US-0139461.
PR 18-JUN-1999;	99US-0139462.
PR 18-JUN-1999;	99US-0139463.
PR 18-JUN-1999;	99US-0139750.
PR 18-JUN-1999;	99US-0139763.
PR 21-JUN-1999;	99US-0139817.
PR 22-JUN-1999;	99US-0139899.
PR 23-JUN-1999;	99US-0140353.
PR 23-JUN-1999;	99US-0140354.
PR 24-JUN-1999;	99US-0140695.
PR 28-JUN-1999;	99US-0140823.
PR 29-JUN-1999;	99US-0140991.
PR 30-JUN-1999;	99US-0141287.
PR 01-JUL-1999;	99US-0141847.
PR 01-JUL-1999;	99US-0142154.
PR 02-JUL-1999;	99US-0142055.
PR 06-JUL-1999;	99US-0142390.
PR 08-JUL-1999;	99US-0142803.
PR 09-JUL-1999;	99US-0142920.
PR 12-JUL-1999;	99US-0142977.
PR 13-JUL-1999;	99US-0143547.
PR 14-JUL-1999;	99US-0143624.
PR 15-JUL-1999;	99US-0144005.
PR 16-JUL-1999;	99US-0144085.
PR 16-JUL-1999;	99US-0144086.
PR 19-JUL-1999;	99US-0144325.
PR 19-JUL-1999;	99US-0144331.
PR 19-JUL-1999;	99US-0144333.
PR 19-JUL-1999;	99US-0144334.
PR 19-JUL-1999;	99US-0144335.
PR 19-JUL-1999;	99US-0144337.
PR 20-JUL-1999;	99US-0144632.
PR 20-JUL-1999;	99US-0144684.
PR 21-JUL-1999;	99US-0144814.
PR 21-JUL-1999;	99US-0145086.
PR 21-JUL-1999;	99US-0145088.
PR 22-JUL-1999;	99US-0145085.

PR 22-JUL-1999;	99US-0145087.
PR 22-JUL-1999;	99US-0145089.
PR 22-JUL-1999;	99US-0145192.
PR 23-JUL-1999;	99US-0145145.
PR 23-JUL-1999;	99US-0145218.
PR 23-JUL-1999;	99US-0145224.
PR 26-JUL-1999;	99US-0145276.
PR 27-JUL-1999;	99US-0145913.
PR 27-JUL-1999;	99US-0145918.
PR 27-JUL-1999;	99US-0145919.
PR 28-JUL-1999;	99US-0145951.
PR 02-AUG-1999;	99US-0146386.
PR 02-AUG-1999;	99US-0146388.
PR 02-AUG-1999;	99US-0146389.
PR 03-AUG-1999;	99US-0147038.
PR 04-AUG-1999;	99US-0147204.
PR 04-AUG-1999;	99US-0147302.
PR 05-AUG-1999;	99US-0147192.
PR 05-AUG-1999;	99US-0147260.
PR 06-AUG-1999;	99US-0147303.
PR 06-AUG-1999;	99US-0147416.
PR 09-AUG-1999;	99US-0147493.
PR 09-AUG-1999;	99US-0147935.
PR 10-AUG-1999;	99US-0148171.
PR 11-AUG-1999;	99US-0148319.
PR 12-AUG-1999;	99US-0148341.
PR 13-AUG-1999;	99US-0148565.
PR 13-AUG-1999;	99US-0148684.
PR 16-AUG-1999;	99US-0149368.
PR 17-AUG-1999;	99US-0149175.
PR 18-AUG-1999;	99US-0149426.
PR 20-AUG-1999;	99US-0149722.
PR 20-AUG-1999;	99US-0149723.
PR 20-AUG-1999;	99US-0149929.
PR 23-AUG-1999;	99US-0149902.
PR 23-AUG-1999;	99US-0149930.
PR 25-AUG-1999;	99US-0150560.
PR 26-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151065.
PR 27-AUG-1999;	99US-0151066.
PR 27-AUG-1999;	99US-0151080.
PR 30-AUG-1999;	99US-0151303.
PR 31-AUG-1999;	99US-0151438.
PR 01-SEP-1999;	99US-0151930.
PR 07-SEP-1999;	99US-0152363.
PR 10-SEP-1999;	99US-0153070.
PR 13-SEP-1999;	99US-0153758.
PR 15-SEP-1999;	99US-0154018.
PR 16-SEP-1999;	99US-0154039.
PR 20-SEP-1999;	99US-0154779.
PR 22-SEP-1999;	99US-0155139.
PR 23-SEP-1999;	99US-0155486.
PR 24-SEP-1999;	99US-0155659.
PR 28-SEP-1999;	99US-0156458.
PR 29-SEP-1999;	99US-0156596.
PR 04-OCT-1999;	99US-0157117.
PR 05-OCT-1999;	99US-0157753.
PR 06-OCT-1999;	99US-0157865.
PR 07-OCT-1999;	99US-0158029.
PR 08-OCT-1999;	99US-0158232.
PR 12-OCT-1999;	99US-0158369.
PR 13-OCT-1999;	99US-0159293.
PR 13-OCT-1999;	99US-0159294.
PR 13-OCT-1999;	99US-0159295.
PR 14-OCT-1999;	99US-0159329.
PR 14-OCT-1999;	99US-0159330.
PR 14-OCT-1999;	99US-0159331.
PR 14-OCT-1999;	99US-0159637.
PR 14-OCT-1999;	99US-0159638.
PR 18-OCT-1999;	99US-0159584.
PR 21-OCT-1999;	99US-0160741.
PR 21-OCT-1999;	99US-0160767.
PR 21-OCT-1999;	99US-0160768.

PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 8; DB 21; Length 278;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 174 SLSEFLDR 181
|||||||
Db 271 SLSEFLDR 278

RESULT 11
AAG1136
ID AAG1136 standard; Protein; 283 AA.
XX AAG1136;
AC AAG1136;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 9740.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139152.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142803.
PR 08-JUL-1999; 99US-0142830.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.

PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 8; DB 21; Length 283;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 SLSFDLDR 181
 |||||
 Db 276 slsfdldr 283

RESULT 12
 AAG11135
 ID AAG11135 standard; Protein; 299 AA.
 XX
 AC AAG11135;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 9739.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.

PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	99US-0136392.	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	99US-0136782.	PR	05-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	PR	06-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137528.	PR	07-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	08-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	09-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	PR	10-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	11-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	12-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	13-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	14-AUG-1999;	99US-0148341.
PR	17-JUN-1999;	99US-0139453.	PR	15-AUG-1999;	99US-0148565.
PR	18-JUN-1999;	99US-0139492.	PR	16-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	17-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	18-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	19-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	21-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	22-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	24-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	28-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	29-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	02-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	03-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	04-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	05-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	06-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	07-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	08-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	09-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	10-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	11-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	12-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	13-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	14-OCT-1999;	99US-0157553.
PR	15-JUL-1999;	99US-0144005.	PR	15-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	16-OCT-1999;	99US-0158029.
PR	19-JUL-1999;	99US-0144325.	PR	17-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144331.	PR	18-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144332.	PR	19-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144333.	PR	20-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144334.	PR	21-OCT-1999;	99US-0159295.
PR	20-JUL-1999;	99US-0144335.	PR	22-OCT-1999;	99US-0159329.
PR	20-JUL-1999;	99US-0144352.	PR	23-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144632.	PR	24-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144884.	PR	25-OCT-1999;	99US-0159637.
PR	21-JUL-1999;	99US-0144814.	PR	26-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0145086.	PR	27-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145088.	PR	28-OCT-1999;	99US-0160741.
PR	22-JUL-1999;	99US-0145085.	PR	29-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145087.	PR	30-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145089.	PR	31-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145192.	PR	01-NOV-1999;	99US-0160814.
PR	23-JUL-1999;	99US-0145145.	PR	02-NOV-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145218.	PR	03-NOV-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145224.	PR	04-NOV-1999;	99US-0160981.
PR	26-JUL-1999;	99US-0145276.	PR	05-NOV-1999;	99US-0160989.
PR	27-JUL-1999;	99US-0145913.	PR	06-NOV-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145918.	PR	07-NOV-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145919.	PR	08-NOV-1999;	99US-0161406.
PR	28-JUL-1999;	99US-0145951.	PR	09-NOV-1999;	99US-0161359.
PR	02-AUG-1999;	99US-0146386.	PR	10-NOV-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146388.	PR	11-NOV-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146389.	PR	12-NOV-1999;	99US-0161920.
PR			PR	13-NOV-1999;	99US-0161922.

PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 8; DB 21; Length 299;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 174 SLSFDLDR 181
|||||||
Db 292 slsfdldr 299

RESULT 13

AAAS2127
ID AAR52127 standard; Peptide; 16 AA.

XX AC AAR52127;

XX DT 27-SEP-1996 (first entry)

XX DE Mouse light chain surface patch S03410.

XX KW antibody; humanised; murine; human; heavy chain; light; variable;
XX KW framework region; complementarity determining region; reshaping;
XX KW modelling; surface residue; modify.

XX OS Mus sp.

XX PN EP592106-A1.

XX PD 13-APR-1994.

XX PF 07-SEP-1993; 93EP-0307051.

XX PR 09-SEP-1992; 92US-0942245.

XX PA (PEDE/) PEDERSEN J T.

XX PA (IMMU-) IMMUNOGEN INC.

XX PI Guild BC, Pedersen JT, Rees AR, Roguska MA, Searle SMJ;

XX DR WPI; 1994-120230/15.

XX PT Method of resurfacing of rodent antibodies to produce humanised
XX PT antibody forms - for producing non-human antibodies with improved
XX PT therapeutic efficiency by presenting human surface on V-region

XX PS Example 1; Page 14; 230pp; English.

XX CC Modification of a rodent antibody or fragment by resurfacing in order
XX CC to produce a humanised rodent antibody can be determined by calculating
XX CC homology between murine and human antibody surfaces. In order to test
XX CC the resurfacing approach of the invention, three humanisation
XX CC experiments were set up. (1) traditional loop grafting; (2) resurfacing
XX CC approach using most similar chain; and (3) resurfacing approach using
XX CC human sequences with most similar surface residues. AAR52069-159 are the
XX CC surface residue patterns in mouse light chain antibody variable regions.
XX CC These "patches" were used in the third method, where rodent light and
XX CC heavy chains were matched and the most similar human sequence found
XX CC independently only over the surface residues indicated in AAR52030-67.

XX SQ Sequence 16 AA;

Query Match 1.6%; Score 7; DB 15; Length 16;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 387 PRSGKGA 393

Db 4 prsgkga 10

RESULT 14

AAV79483

XX AC AAV79483 standard; peptide; 50 AA.

XX AC AAV79483;

XX DT 01-AUG-2000 (first entry)

XX DE Rat alpha tubulin variant peptide (aa314-363).

XX KW Rat; alpha tubulin; mutant; mutein; variant;
XX KW metabolic x syndrome; hypertension; stroke; diabetes;
XX KW insulin resistance; obesity; dyslipidemia; anorectic;
XX KW hypotensive; cerebroprotective; vasotropic; antidiabetic;
XX KW antilipemic; gene therapy; spontaneously hypertensive; SHR.

XX OS Rattus sp.

XX FH Key Location/Qualifiers

XX FT Peptide 1..50

XX FT /note= "corresponds to residues 314-363 of
XX FT full-length protein"

XX FT Misc-difference 27

XX FT /note= "replaces Thr in wild-type"

XX PN WO200018918-A2.

XX PD 06-APR-2000.

XX PF 28-SEP-1999; 99WO-US22494.

XX PR 28-SEP-1998; 98US-0161939.

XX PA (CURA-) CURAGEN CORP.

XX PI Shimkets RA;

XX DR WPI; 2000-303450/26.

XX PT Novel genes encoding e.g. CD36, SGLT2, and kynurenine aminotransferase,
XX PT which are predictive and therapeutic for stroke, hypertension, diabetes
XX PT and obesity

XX PS Claim 2; Fig 4G; 79pp; English.

XX CC This sequence represents amino acids 314-363 of alpha-tubulin
XX CC from a spontaneously hypertensive rat (SHR). This region of
XX CC alpha-tubulin carries a Ser for Thr amino acid substitution when
XX CC compared to the corresponding region in control Wistar Kyoto (WKY)
XX CC rats (see AAV79484). The alpha-tubulin gene is differentially
XX CC expressed in SHR, SHR-stroke prone (SP) and WKY rats. It may be
XX CC involved in a predisposition to vascular injury. The SHR rat
XX CC is an animal model for human metabolic X syndrome. Genes encoding
XX CC sodium dependent glucose cotransporter, kynurenine aminotransferase,
XX CC CD36, aldolase A, atrial natriuretic factor, alpha-cardiac myosin
XX CC and alpha-tubulin were identified as being potentially associated
XX CC with hypertension, obesity and insulin resistance. These
XX CC proteins, including isolated proteins comprising the present
XX CC sequence, are used for treating, preventing and diagnosing
XX CC ischemic and metabolic diseases and disorders, such as stroke,
XX CC hypertension, diabetes and obesity, especially insulin resistivity,
XX CC dyslipidemia and ischemic stroke (all claimed). Polynucleotides
XX CC encoding them may also be used in gene therapy and antisense
XX CC therapy protocols.

XX SQ Sequence 50 AA;

Query Match 1.6%; Score 7; DB 21; Length 50;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 114 IKTKRSI 120

```
Db      22 iktkrs1 28
|||||||
RESULT 15
AAG06322
ID      AAG06322 standard; Protein; 64 AA.
XX
AC      AAG06322;
XX
DT      17-OCT-2000 (first entry)
XX
DE      Arabidopsis thaliana protein fragment SEQ ID NO: 3055.
XX
KW      Protein identification; signal transduction pathway; metabolic pathway;
KW      hybridisation assay; genetic mapping; gene expression control; promoter;
KW      termination sequence.
XX
OS      Arabidopsis thaliana.
PN      EP1033405-A2.
XX
PD      06-SEP-2000.
XX
PF      25-FEB-2000; 2000EP-0301439.
XX
PR      25-FEB-1999; 99US-0121825.
PR      05-MAR-1999; 99US-0123180.
PR      09-MAR-1999; 99US-0123548.
PR      23-MAR-1999; 99US-0125788.
PR      25-MAR-1999; 99US-0126264.
PR      29-MAR-1999; 99US-0126785.
PR      01-APR-1999; 99US-0127462.
PR      06-APR-1999; 99US-0128234.
PR      08-APR-1999; 99US-0128714.
PR      16-APR-1999; 99US-0129845.
PR      19-APR-1999; 99US-0130077.
PR      21-APR-1999; 99US-0130449.
PR      23-APR-1999; 99US-0130510.
PR      28-APR-1999; 99US-0130891.
PR      30-APR-1999; 99US-0131449.
PR      04-MAY-1999; 99US-0132407.
PR      05-MAY-1999; 99US-0132485.
PR      06-MAY-1999; 99US-0132486.
PR      07-MAY-1999; 99US-0132863.
PR      11-MAY-1999; 99US-0134256.
PR      14-MAY-1999; 99US-0134218.
PR      14-MAY-1999; 99US-0134219.
PR      14-MAY-1999; 99US-0134221.
PR      18-MAY-1999; 99US-0134370.
PR      19-MAY-1999; 99US-0134941.
PR      20-MAY-1999; 99US-0135124.
PR      21-MAY-1999; 99US-0135353.
PR      24-MAY-1999; 99US-0135629.
PR      25-MAY-1999; 99US-0136021.
PR      27-MAY-1999; 99US-0136392.
PR      28-MAY-1999; 99US-0136782.
PR      01-JUN-1999; 99US-0137222.
PR      03-JUN-1999; 99US-0137528.
PR      04-JUN-1999; 99US-0137502.
PR      07-JUN-1999; 99US-0137724.
PR      08-JUN-1999; 99US-0138094.
PR      10-JUN-1999; 99US-0138540.
PR      10-JUN-1999; 99US-0138847.
PR      14-JUN-1999; 99US-0139119.
PR      16-JUN-1999; 99US-0139452.
PR      16-JUN-1999; 99US-0139453.
PR      17-JUN-1999; 99US-0139492.
PR      18-JUN-1999; 99US-0139454.
PR      18-JUN-1999; 99US-0139455.
PR      18-JUN-1999; 99US-0139456.
PR      18-JUN-1999; 99US-0139457.
PR      18-JUN-1999; 99US-0139458.
PR      18-JUN-1999; 99US-0139459.
PR      18-JUN-1999; 99US-0139460.
PR      18-JUN-1999; 99US-0139461.
PR      18-JUN-1999; 99US-0139462.
PR      18-JUN-1999; 99US-0139463.
PR      18-JUN-1999; 99US-0139750.
PR      18-JUN-1999; 99US-0139763.
PR      21-JUN-1999; 99US-0139817.
PR      22-JUN-1999; 99US-0139899.
PR      23-JUN-1999; 99US-0140353.
PR      23-JUN-1999; 99US-0140354.
PR      24-JUN-1999; 99US-0140695.
PR      28-JUN-1999; 99US-0140823.
PR      29-JUN-1999; 99US-0140991.
PR      30-JUN-1999; 99US-0141287.
PR      01-JUL-1999; 99US-0141842.
PR      01-JUL-1999; 99US-0142154.
PR      02-JUL-1999; 99US-0142055.
PR      06-JUL-1999; 99US-0142390.
PR      08-JUL-1999; 99US-0142803.
PR      09-JUL-1999; 99US-0142920.
PR      12-JUL-1999; 99US-0142977.
PR      13-JUL-1999; 99US-0143542.
PR      14-JUL-1999; 99US-0143624.
PR      15-JUL-1999; 99US-0144005.
PR      16-JUL-1999; 99US-0144085.
PR      16-JUL-1999; 99US-0144086.
PR      19-JUL-1999; 99US-0144325.
PR      19-JUL-1999; 99US-0144331.
PR      19-JUL-1999; 99US-0144332.
PR      19-JUL-1999; 99US-0144334.
PR      19-JUL-1999; 99US-0144335.
PR      20-JUL-1999; 99US-0144352.
PR      20-JUL-1999; 99US-0144632.
PR      20-JUL-1999; 99US-0144884.
PR      21-JUL-1999; 99US-0144814.
PR      21-JUL-1999; 99US-0145086.
PR      21-JUL-1999; 99US-0145088.
PR      22-JUL-1999; 99US-0145085.
PR      22-JUL-1999; 99US-0145087.
PR      22-JUL-1999; 99US-0145089.
PR      22-JUL-1999; 99US-0145192.
PR      23-JUL-1999; 99US-0145145.
PR      23-JUL-1999; 99US-0145218.
PR      23-JUL-1999; 99US-0145224.
PR      26-JUL-1999; 99US-0145276.
PR      27-JUL-1999; 99US-0145913.
PR      27-JUL-1999; 99US-0145918.
PR      27-JUL-1999; 99US-0145919.
PR      28-JUL-1999; 99US-0145951.
PR      02-AUG-1999; 99US-0146386.
PR      02-AUG-1999; 99US-0146388.
PR      02-AUG-1999; 99US-0146389.
PR      03-AUG-1999; 99US-0147038.
PR      04-AUG-1999; 99US-0147204.
PR      04-AUG-1999; 99US-0147302.
PR      05-AUG-1999; 99US-0147192.
PR      05-AUG-1999; 99US-0147260.
PR      06-AUG-1999; 99US-0147303.
PR      06-AUG-1999; 99US-0147416.
PR      09-AUG-1999; 99US-0147493.
PR      09-AUG-1999; 99US-0147935.
PR      10-AUG-1999; 99US-0148171.
PR      11-AUG-1999; 99US-0148319.
PR      12-AUG-1999; 99US-0148341.
PR      13-AUG-1999; 99US-0148565.
PR      13-AUG-1999; 99US-0148684.
PR      16-AUG-1999; 99US-0149368.
```

```
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.6%; Score 7; DB 21; Length 64;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 162 QLEYAAK 168
Db 44 qleyaak 50
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:46:30 ; Search time 36.87 Seconds
(without alignments)
269.771 Million cell updates/sec

Title: US-09-787-083-2
Perfect score: 442
Sequence: 1 MKVSLSTLTLSCFAILA.....YNHEATSPGVGLMLNDWGL 442

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 212252 seqs, 22503292 residues

Word size : 0
Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	DB ID	Description
1	7	1.6	16	US-07-942-245-98
2	7	1.6	347	Sequence 98, Appl
3	7	1.6	552	Sequence 14, Appl
4	7	1.6	900	Sequence 6, Appl
5	7	1.6	934	Sequence 4, Appl
6	7	1.6	1026	Sequence 80, Appl
7	7	1.6	1026	Sequence 6, Appl
8	6	1.4	10	Sequence 6, Appl
9	6	1.4	11	Sequence 30, Appl
10	6	1.4	12	Sequence 29, Appl
11	6	1.4	13	Sequence 28, Appl
12	6	1.4	13	Sequence 27, Appl
13	6	1.4	14	Sequence 27, Appl
14	6	1.4	14	Sequence 10, Appl
15	6	1.4	14	Sequence 32, Appl
16	6	1.4	15	Sequence 5, Appl
17	6	1.4	16	Sequence 4, Appl
18	6	1.4	20	Sequence 1, Appl
19	6	1.4	21	Sequence 13, Appl
20	6	1.4	21	Sequence 15, Appl
21	6	1.4	21	Sequence 1, Appl
22	6	1.4	21	Sequence 1, Appl
23	6	1.4	21	Sequence 13, Appl
24	6	1.4	21	Sequence 15, Appl
25	6	1.4	21	Sequence 1, Appl
26	6	1.4	21	Sequence 1, Appl
27	6	1.4	21	Sequence 1, Appl

28	6	1.4	21	1	US-08-285-661-1	Sequence 1, Appl
29	6	1.4	21	1	US-08-301-838-1	Sequence 1, Appl
30	6	1.4	21	1	US-08-389-487-4	Sequence 4, Appl
31	6	1.4	21	1	US-08-389-487-12	Sequence 12, Appl
32	6	1.4	21	1	US-08-507-124-2	Sequence 2, Appl
33	6	1.4	21	1	US-08-507-124-3	Sequence 3, Appl
34	6	1.4	21	1	US-08-342-931-1	Sequence 1, Appl
35	6	1.4	21	1	US-08-400-256-1	Sequence 1, Appl
36	6	1.4	21	2	US-08-508-664-9	Sequence 85, Appl
37	6	1.4	21	2	US-08-353-476-85	Sequence 87, Appl
38	6	1.4	21	2	US-08-484-219-1	Sequence 1, Appl
39	6	1.4	21	2	US-08-379-587-1	Sequence 1, Appl
40	6	1.4	21	2	US-08-992-676-1	Sequence 1, Appl
41	6	1.4	21	2	US-08-992-676-6	Sequence 6, Appl
42	6	1.4	21	2	US-09-134-836-1	Sequence 1, Appl
43	6	1.4	21	2	US-09-134-836-7	Sequence 7, Appl
44	6	1.4	21	2	US-08-932-082-1	Sequence 1, Appl
45	6	1.4	21	4	US-08-932-082-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-07-942-245-98
; Sequence 98, Application US/07942245
; Patent No. 5639641
; GENERAL INFORMATION:
; APPLICANT: PEDERSEN, Jan T.
; APPLICANT: SEARLE, Stephen M.J.
; APPLICANT: REES, Anthony R.
; APPLICANT: ROGUSKA, Michael A.
; APPLICANT: GUILD, Braydon C.
; TITLE OF INVENTION: SURFACE RESIDUE VENEERING OF RODENT
; TITLE OF INVENTION: ANTIBODIES
; NUMBER OF SEQUENCES: 522
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: UNIX
; SOFTWARE: In house
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/942,245
; FILING DATE: 09-SEP-1992
; CLASSIFICATION: 530
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-942-245-98

Query Match 1.6%; Score 7; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 387 PRSGKA 393
Db 4 PRSGKA 10

```
RESULT 2
US-09-889-14
; Sequence 14, Application US/09097889
; Patent No. 6218117
; GENERAL INFORMATION:
; APPLICANT: HerinStadt, Corrina
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Davis, Robert E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
; TITLE OF INVENTION: AGENTS THAT QUANTITATIVELY ALTER DETECTABLE
; TITLE OF INVENTION: EXTRAMITOCHONDRIAL DNA: MITOCHONDRIAL DNA RATIOS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/097,889
; FILING DATE: 15-JUN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rosenman Ph.D., Stephen J.
; REGISTRATION NUMBER: 43,058
; REFERENCE/DOCKET NUMBER: 660088.417
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-09-097-889-14

Query Match 1.6%; Score 7; DB 4; Length 347;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 L7LSILS 14
Db 155 L7LSILS 161

RESULT 3
US-08-231-729B-6
; Sequence 6, Application US/08231729B
; Patent No. 5618722
; GENERAL INFORMATION:
; APPLICANT: ZENNO, Shuhei
; APPLICANT: SHIRAIISHI, Shinji
; APPLICANT: INOUE, Satoshi
; APPLICANT: SAIGO, Kaoru
; TITLE OF INVENTION: FIREFLY LUCIFERASE GENE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LEYDIG, VOIT & MAYER
; STREET: 700 Thirteenth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/890,865A
; FILING DATE: 10-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/54249
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)278-0400
; TELEFAX: (212)391-0526
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 900 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,729B
; FILING DATE: 20-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 119050/1993
; FILING DATE: 21-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rose, Herbert C.
; REGISTRATION NUMBER: 29846
; REFERENCE/DOCKET NUMBER: 60130/No. 5618722aka
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-737-6770
; TELEFAX: 202-737-6776
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 552 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-231-729B-6

Query Match 1.6%; Score 7; DB 1; Length 552;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 LSFDLDR 181
Db 184 LSFDLDR 190

RESULT 4
US-08-890-865A-4
; Sequence 4, Application US/08890865A
; Patent No. 6307019
; GENERAL INFORMATION:
; APPLICANT: Constantini, Franklin
; APPLICANT: Zeng, Li
; TITLE OF INVENTION: AXIN GENE AND USES THEREOF
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/890,865A
; FILING DATE: 10-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/54249
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)278-0400
; TELEFAX: (212)391-0526
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 900 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-890-865A-4

Query Match 1.6%; Score 7; DB 4; Length 900;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 403 KGISGY 409
| | | | |
Db 268 KGISGY 274

RESULT 5

US-08-215-805A-80
; Sequence 80, Application US/08215805A
; Patent No. 5559008
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; TITLE OF INVENTION: LEUKOTOXIN GENE FROM PASTEURRELLA
; NUMBER OF SEQUENCES: 84
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/215,805A
; FILING DATE: 22-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Timian, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 19603/61 (D-1329A)
; TELEPHONE: (716) 263-1636
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 80:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 934 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Pasteurella suis
; STRAIN: 5943
; IMMEDIATE SOURCE:
; LIBRARY: P. suis DNA in Bacteriophage lambda-dash
; CLONE: (Lambda)yfc33-37
US-08-215-805A-80

Query Match 1.6%; Score 7; DB 1; Length 934;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 238 VKAAEDL 244
| | | | |
Db 62 VKAAEDL 68

RESULT 6

US-08-542-003-6

; Sequence 6, Application US/08542003
; Patent No. 5864013
; GENERAL INFORMATION:
; APPLICANT: Goldberg, Edward B.
; TITLE OF INVENTION: MATERIALS FOR THE PRODUCTION OF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie and Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/542,003
; FILING DATE: 13-OCT-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 8471-0005-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: 212-869-8864
; TELEX: 66441 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1026 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Bacteriophage T4
; IMMEDIATE SOURCE:
; CLONE: p37 amino acid
US-08-542-003-6

Query Match 1.6%; Score 7; DB 2; Length 1026;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 QTKGKP 134
| | | | |
Db 861 QTKGKP 867

RESULT 7

US-08-322-760A-6
; Sequence 6, Application US/08322760A
; Patent No. 5877279
; GENERAL INFORMATION:
; APPLICANT: Goldberg, Edward B.
; TITLE OF INVENTION: MATERIALS FOR THE PRODUCTION OF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie and Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/322,760A
FILING DATE: 13-OCT-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 8471-0003-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: 212-869-8864
TELEX: 66441 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1026 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Bacteriophage T4
IMMEDIATE SOURCE:
CLONE: p37 amino acid
US-08-322-760A-6

Query Match 1.6%; Score 7; DB 2; Length 1026;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 QTIKGP 134
|||||||
DB 861 QTIKGP 867

RESULT 8
US-08-974-775-30
Sequence 30, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:
APPLICANT: Toback, F. Gary
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775
FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid

STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-30

Query Match 1.4%; Score 6; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
|||||||
DB 5 NHEATS 10

RESULT 9
US-08-974-775-29
Sequence 29, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:
APPLICANT: Toback, F. Gary
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775
FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-29

Query Match 1.4%; Score 6; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
|||||||
DB 5 NHEATS 10

RESULT 10
US-08-974-775-28
Sequence 28, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:

```
;
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-974-775-28

Query Match 1.4%; Score 6; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 424 NHEATS 429
Db 5 NHEATS 10

RESULT 11
US-08-974-775-8
; Sequence 8, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
```

```
;
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-974-775-8

Query Match 1.4%; Score 6; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 424 NHEATS 429
Db 8 NHEATS 13

RESULT 12
US-08-974-775-27
; Sequence 27, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-974-775-27

Query Match 1.4%; Score 6; DB 3; Length 13;
```

Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

QY 424 NHEATS 429
|||||
Db 5 NHEATS 10

RESULT 13

US-08-974-775-9
; Sequence 9, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299

INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-974-775-9

Query Match 1.4%; Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

QY 424 NHEATS 429
|||||
Db 8 NHEATS 13

RESULT 14

US-08-974-775-10
; Sequence 10, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299

INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-974-775-9

Query Match 1.4%; Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

QY 424 NHEATS 429
|||||
Db 8 NHEATS 13

RESULT 14

US-08-974-775-10
; Sequence 10, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:

STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; STREET: Plaza Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299

INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-974-775-10

Query Match 1.4%; Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

QY 424 NHEATS 429
|||||
Db 8 NHEATS 13

RESULT 15

US-08-974-775-32
; Sequence 32, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299

INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-974-775-10

Query Match 1.4%; Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

QY 424 NHEATS 429
|||||
Db 8 NHEATS 13

RESULT 15

US-08-974-775-32
; Sequence 32, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:

Query Match 1.4%; Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

QY 424 NHEATS 429
|||||
Db 8 NHEATS 13

RESULT 15

US-08-974-775-32
; Sequence 32, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:

Query Match 1.4%; Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0;

Indels 0; Gaps 0;

QY 424 NHEATS 429
|||||
Db 8 NHEATS 13

RESULT 15

US-08-974-775-32
; Sequence 32, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599

```

; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 14
; OTHER INFORMATION: /product= "Ser or Tyr"
US-08-974-775-32

```

```

Query Match      1.4%; Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
    |||||
Db 8 NHEATS 13

```

Search completed: May 9, 2002, 12:55:41
Job time: 551 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:46:30 ; Search time 44.37 Seconds
(without alignments)
758.827 Million cell updates/sec

Title: US-09-787-083-2

Perfect score: 442

Sequence: 1 MKVSLSTLTLSILSCFAIIA.....YNHEATSFVGLMLNDWMGL 442

Scoring table:

OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

PIR_68.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	9	2.0	382	2	E81195	phospholipase A1,
2	9	2.0	409	2	H81831	probable phospholip
3	8	1.8	286	2	B36371	outer membrane pho
4	8	1.8	299	2	T47989	RAV-like protein -
5	8	1.8	1396	2	S36851	L-shaped tail fibe
6	7	1.6	100	2	S61050	probable membrane
7	7	1.6	122	2	E75377	hypothetical prote
8	7	1.6	158	2	E86498	pts IIA Protein [1
9	7	1.6	158	2	E72124	pts iia protein -
10	7	1.6	197	2	A54506	tubulin alpha-1 ch
11	7	1.6	219	2	T19438	hypothetical prote
12	7	1.6	240	2	A61544	tubulin alpha chai
13	7	1.6	264	2	A44959	coat protein - pot
14	7	1.6	267	2	A60366	coat protein - pot
15	7	1.6	267	2	S14001	genome polypeptid
16	7	1.6	267	2	S26630	capsid protein - p
17	7	1.6	267	2	JC1527	coat protein - pot
18	7	1.6	267	2	S13239	tail fiber protein
19	7	1.6	267	2	E84092	phage-related prot
20	7	1.6	270	2	S73734	abc transport ATP-
21	7	1.6	273	2	C72328	transaminase B hom
22	7	1.6	287	2	A72421	oligopeptide ABC t
23	7	1.6	288	2	T36237	probable phosphotr
24	7	1.6	306	2	T27985	hypothetical prote
25	7	1.6	310	2	T02200	probable RAV-like
26	7	1.6	313	2	JT0960	polyprotein - pota
27	7	1.6	321	2	H82240	conserved hypothet
28	7	1.6	324	2	T23876	hypothetical prote
29	7	1.6	326	2	A43939	proteinase inhibit

30	7	1.6	338	2	F69437	hypothetical prote
31	7	1.6	341	2	T47653	pectate lyase-like
32	7	1.6	342	2	G64411	hypothetical prote
33	7	1.6	342	2	C64394	hypothetical prote
34	7	1.6	347	1	DNHUN2	NADH dehydrogenase
35	7	1.6	347	2	T11051	NADH dehydrogenase
36	7	1.6	347	2	B59153	NADH dehydrogenase
37	7	1.6	349	2	T43920	yfuc protein [impo
38	7	1.6	370	2	G83219	probable dihydroli
39	7	1.6	379	2	S13556	genome polypeptid
40	7	1.6	381	2	G82132	ribonuclease D vcl
41	7	1.6	382	2	S13237	tail fiber protein
42	7	1.6	384	2	S70638	tubulin alpha chai
43	7	1.6	384	2	S70641	tubulin alpha chai
44	7	1.6	384	2	S70640	tubulin alpha chai
45	7	1.6	387	2	S33517	tubulin alpha chai

ALIGNMENTS

RESULT 1

E81195

phospholipase A1, probable NMB0464 [imported] - Neisseria meningitidis (strain MC58 s
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: E81195
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755
A;Accession: E81195
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-382 <TET>
A;Cross-references: GB:AE002403; GB:AE002098; NID:g7225688; PIDN:AAF40901.1; PID:g722
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0464

Query Match 2.0% Score 9; DB 2; Length 382;
Best Local Similarity 100.0%; Pred. No. 0.45; Mismatches 0; Indels 0; Caps 0;
Matches 9; Conservative 0;

Qy 279 PEIFLTQPV 287
|||||||
Db 221 PEIFLTQPV 229

RESULT 2

H81831

probable phospholipase NMA2021 [imported] - Neisseria meningitidis (strain Z2491 serog
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C;Accession: H81831
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo
; Holroyd, S.; Jagels, K.; Leather, S.; Mungall, K.; Quail, M.A.; Rajandre
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491
A;Reference number: A81775; MUID:20222556
A;Accession: H81831
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-409 <PAR>
A;Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85240.1; PID:g738
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA2021

Query Match
Best Local Similarity 2.0%; Score 9; DB 2; Length 409;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIFLTQPV 287
|||||
DB 248 PEIFLTQPV 256

RESULT 3
B36971
outer membrane phospholipase A (EC 3.1.1.-) precursor - Klebsiella pneumoniae
C:Species: Klebsiella pneumoniae
C>Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 18-Jun-1999
C:Accession: B36971; S40129
R:Brok, R.G.P.M.; Brinkman, E.; van Bortel, R.; Bekkers, A.C.A.P.A.; Verheij, H.M.; Tomm
J. Bacteriol. 176, 861-870, 1994
A:Title: Molecular characterization of enterobacterial plda genes encoding outer membran
A:Reference number: A36971; MUID:94131966
A:Accession: B36971
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-286 <BRO>
A:Cross-references: EMBL:X76901; NID:g436880; PIDN:CAA54223.1; PID:g436881
A:Note: authors translated the codon AAG for residue 112 as Arg
C:Genetics:
A:Gene: plda
C:Superfamily: bacterial phospholipase A1
C:Keywords: carboxylic ester hydrolase

Query Match
Best Local Similarity 1.8%; Score 8; DB 2; Length 286;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 431 GVGGLMLND 438
|||||
DB 277 GVGGLMLND 284

RESULT 4
T47989
RAV-like protein - Arabidopsis thaliana
N:Alternate names: protein F21F14.140
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47989
R:Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Sa
submitted to the Protein Sequence Database, February 2000
A:Reference number: 224481
A:Accession: T47989
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-299 <CHO>
A:Cross-references: EMBL:AL138642
A:Experimental source: cultivar Columbia; BAC clone F21F14
C:Genetics:
A:Map position: 3
A:Note: F21F14.140

Query Match
Best Local Similarity 1.8%; Score 8; DB 2; Length 299;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 SLSFDLDR 181
|||||
DB 292 SLSFDLDR 299

RESULT 5
S36851

L-shaped tail fiber protein - phage T5
N:Alternate names: ltf protein
C:Species: phage T5
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 17-Nov-2000
C:Accession: S65934; S01984; S36851
R:Kallman, A.V.; Kulshin, V.E.; Shlyapnikov, M.G.; Ksenzenko, V.N.; Kryukov, V.M.
FEBS Lett. 366, 46-48, 1995
A:Title: The nucleotide sequence of the bacteriophage T5 ltf gene.
A:Reference number: S65934; MUID:95309401
A:Accession: S65934
A:Molecule type: DNA
A:Residues: 1-1396 <KAL>
A:Cross-references: EMBL:X69460; NID:g15415; PIDN:CAA49220.1; PID:g15416
R:Kallman, A.V.; Kryukov, V.M.; Bayev, A.A.
Nucleic Acids Res. 16, 6230, 1988
A:Title: The nucleotide sequence of bacteriophage T5 DNA at the region between early
A:Reference number: S01982; MUID:86289370
A:Accession: S01984
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 934-985, 'A' 987-1396 <KA2>
A:Cross-references: EMBL:X07559
C:Genetics:
A:Gene: ltf
C:Keywords: late protein; tail fiber

Query Match
Best Local Similarity 1.8%; Score 8; DB 2; Length 1396;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 QSATQSAS 61
|||||
DB 87 QSATQSAS 94

RESULT 6
S61050
probable membrane protein YDL163w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein D1505
C:Species: Saccharomyces cerevisiae
C>Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 20-Jun-2000
C:Accession: S61050; S67715
R:Pohl, T.M.
submitted to the EMBL Data Library, November 1995
A:Reference number: S61010
A:Accession: S61050
A:Molecule type: DNA
A:Residues: 1-100 <POH>
A:Cross-references: EMBL:Z67750; NID:g1061256; PID:g1061276
R:Pohl, T.M.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67708
A:Accession: S67715
A:Molecule type: DNA
A:Residues: 1-100 <POW>
A:Cross-references: EMBL:Z74212; NID:g1431258; PID:g1431260; MIPS:YDL163w
A:Experimental source: strain S288C
C:Genetics:
A:Map position: 4L
C:Superfamily: Saccharomyces probable membrane protein YDL163w
C:Keywords: transmembrane protein
F:3-19/Domain: transmembrane #status predicted <TMM>

Query Match
Best Local Similarity 1.6%; Score 7; DB 2; Length 100;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 LTLSILS 14
|||||
DB 26 LTLSILS 32

RESULT 7

E75377
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: E75377
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: E75377
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-122 <WHI>
A:Cross-references: GB:AE002002; GB:AE000513; NID:g6459345; PIDN:AAF11150.1; PID:g6459345
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1580
A:Map position: 1

Query Match

Best Local Similarity 1.6%; Score 7; DB 2; Length 122;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 70 EHEPELY 76

|||||

Db 115 EHEPELY 121

RESULT 8

E86498
Pts IIA Protein [imported] - Chlamydomophila pneumoniae (strain J138)
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: E86498
R:Shiral, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349
A:Accession: E86498
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-158 <STO>
A:Cross-references: GB:BA000008; NID:g8978434; PIDN:BAA98271.1; GSPDB:GN00142
A:Experimental source: strain J138
C:Genetics:
A:Gene: ptsN_1
C:Superfamily: phosphotransferase system enzyme II; phosphotransferase system mannitol-s

Query Match

Best Local Similarity 1.6%; Score 7; DB 2; Length 158;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 LSTLTLS 11

|||||

Db 126 LSTLTLS 132

RESULT 9

E72124
pts iia protein - Chlamydomophila pneumoniae (strain CWL029)
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
C:Accession: E72124
R:Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606

Query Match

Best Local Similarity 1.6%; Score 7; DB 2; Length 219;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

A:Accession: E72124
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-158 <ARN>
A:Cross-references: GB:AE001591; GB:AE001363; NID:g4376311; PIDN:AAD18213.1; PID:g437
A:Experimental source: strain CWL029
C:Genetics:
A:Gene: ptsN_1
C:Superfamily: phosphotransferase system enzyme II; phosphotransferase system mannico
F;28-156/Domain: phosphotransferase system mannitol-specific enzyme II factor III hom

Query Match

Best Local Similarity 1.6%; Score 7; DB 2; Length 158;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 LSTLTLS 11

|||||

Db 126 LSTLTLS 132

RESULT 10

A54506
tubulin alpha-1 chain - Plasmodium yoelii (fragment)
C:Species: Plasmodium yoelii
C:Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 13-Aug-1999
C:Accession: A54506
R:Kellia, R.; Arasu, P.; Vaidya, A.B.
Mol. Biochem. Parasitol. 30, 165-174, 1988
A:Title: Molecular clones of alpha-tubulin genes of Plasmodium yoelii reveal an unusu
A:Reference number: A54506; MUID:89014607
A:Accession: A54506
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-197 <AKE>
A:Cross-references: GB:M29816; NID:g160729; PIDN:AAA29779.1; PID:g160730
C:Superfamily: tubulin

Query Match

Best Local Similarity 1.6%; Score 7; DB 2; Length 197;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 114 IKTKRSI 120

|||||

Db 82 IKTKRSI 88

RESULT 11

T19438
hypothetical protein C25A1.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T19438
R:Mortimore, B.
Submitted to the EMBL Data Library, October 1996
A:Reference number: Z19124
A:Accession: T19438
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-219 <WIL>
A:Cross-references: EMBL:Z81038; PIDN:CAB02762.1; GSPDB:GN000019; CESP:C25A1.1
A:Experimental source: clone C25A1
C:Genetics:
A:Gene: CESP:C25A1.1
A:Map position: 1
A:Introns: 33/1; 66/1; 151/2

Query Match

Best Local Similarity 1.6%; Score 7; DB 2; Length 219;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 158 KDAKQLE 164
|||||||
Db 52 KDAKQLE 58

RESULT 12

A61544
tubulin alpha chain - common octopus (fragment)
C:Species: Octopus vulgaris (common octopus)
C:Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 10-Jul-1998
C:Accession: A61544
R:Zinov'eva, R.D.; Aleinikova, K.S.; Tomarev, S.I.
Dokl. Akad. Nauk SSSR 302, 462-467, 1988
A:Title: Isolation and structural characterization of cDNAs coding for alpha-tubulin of
A:Reference number: A61544
A:Accession: A61544
A>Status: preliminary
A:Molecule type: mRNA
C:Residues: 1-240 <ZIN>
C:Superfamily: tubulin
F:239-240/Cleavage site: Glu-Tyr (tubulin-specific carboxypeptidase) #status predicted
F:239-240/Cross-link: peptide (Glu-Tyr) (by tubulin-tyrosine ligase) #status predicted

Query Match 1.6%; Score 7; DB 2; Length 240;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTKRSI 120
|||||||
Db 124 IKTKRSI 130

RESULT 13

A44959
coat protein - potato virus Y (strain N) (fragment)
C:Species: potato virus Y, PVY
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Nov-2000
C:Accession: A44959
R:Hay, J.M.; Fellowes, A.P.; Timmerman, G.M.
Arch. Virol. 107, 111-122, 1989
A:Title: Nucleotide sequence of the coat protein gene of a necrotic strain of potato virus Y
A:Reference number: A44959; MUID:90025790
A:Accession: A44959
A>Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-264 <HAY>
A:Cross-references: GB:M22470; NID:g333508; PIDN:AAA47182.1; PID:g333509
C:Superfamily: tobacco etch virus genome polyprotein

Query Match 1.6%; Score 7; DB 2; Length 264;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 156 TKKDAKQ 162
|||||||
Db 8 TKKDAKQ 14

RESULT 14

A60366
coat protein - potato virus Y (strain TH)
C:Species: potato virus Y, PVY
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Nov-2000
C:Accession: A60366
R:Hataya, T.; Sano, T.; Ohshima, K.; Shikata, E.
Virus Genes 4, 339-350, 1990
A:Title: Polymerase chain-reaction-mediated cloning and expression of the coat protein gene
A:Reference number: A60366; MUID:91143125
A:Accession: A60366
A>Status: preliminary
A:Molecule type: DNA

A:Residues: 1-267 <HAT>
C:Superfamily: tobacco etch virus genome polyprotein

Query Match 1.6%; Score 7; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 156 TKKDAKQ 162
|||||||
Db 11 TKKDAKQ 17

RESULT 15

S14001
genome polyprotein - potato virus Y (strain H) (fragment)
N:Contains: coat protein
C:Species: potato virus Y, PVY
A:Variety: strain H
C:Date: 21-Nov-1993 #sequence_revision 08-Nov-1996 #text_change 17-Nov-2000
C:Accession: S14001
R:Dalmay, T.; Balazs, E.
Nucleic Acids Res. 18, 6721, 1990
A:Title: Nucleotide sequence of an altered virulence potato virus Y coat protein gene
A:Reference number: S14001; MUID:91067494
A:Accession: S14001
A>Status: translation not shown
A:Molecule type: genomic RNA
A:Residues: 1-267 <DAL>
A:Cross-references: EMBL:X54611; NID:g61449; PIDN:CAA38432.1; PID:g930257
A:Experimental source: strain H
C:Superfamily: tobacco etch virus genome polyprotein
C:Keywords: coat protein; polyprotein
F:1-267/Product: coat protein #status predicted <CPR>

Query Match 1.6%; Score 7; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 156 TKKDAKQ 162
|||||||
Db 11 TKKDAKQ 17

Search completed: May 9, 2002, 12:56:32
Job time: 602 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:55:45 ; Search time 28.06 Seconds
(without alignments)
577.543 Million cell updates/sec

Title: US-09-787-083-2

Perfect score: 442

Sequence: 1 MKVSLSTLTLSLSCFAILA.....YNHEATSGVGLMLNDWML 442

Scoring table:

OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 100059 seqs, 36664827 residues

Word size : 0

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	8	1.8	286	1	PAL_KLEPN
2	8	1.8	1396	1	VLRF_BPT5
3	7	1.6	186	1	PRL_ICTPU
4	7	1.6	197	1	TBA_PLAYO
5	7	1.6	240	1	TBA_OCTVU
6	7	1.6	270	1	Y304_MYCPN
7	7	1.6	273	1	ILVE_THEMA
8	7	1.6	322	1	MIAA_PSEPU
9	7	1.6	326	1	SPI_BACBR
10	7	1.6	342	1	FLAD_METJA
11	7	1.6	342	1	Y755_METJA
12	7	1.6	347	1	NU2M_CAPHI
13	7	1.6	347	1	NU2M_HUMAN
14	7	1.6	347	1	NU2M_PANTR
15	7	1.6	347	1	NU2M_PONPA
16	7	1.6	347	1	NU2M_SHEEP
17	7	1.6	423	1	AMT_AQUAE
18	7	1.6	444	1	TBA_ONCKE
19	7	1.6	448	1	TBA4_HUMAN
20	7	1.6	448	1	TBA5_CHICK
21	7	1.6	449	1	TBA2_DROME
22	7	1.6	449	1	TBAE_PHYPO
23	7	1.6	449	1	TBAN_PHYPO
24	7	1.6	449	1	TBA_XENLA
25	7	1.6	450	1	TBAD_PHYPO
26	7	1.6	450	1	TBA_HAECO
27	7	1.6	450	1	TBA_NOTVI
28	7	1.6	451	1	TBAL_CRIGR
29	7	1.6	451	1	TBA3_HOMAM
30	7	1.6	451	1	TBA_OCTDO
31	7	1.6	451	1	TBA_TORNA
32	7	1.6	511	1	LGT_CITUN
33	7	1.6	513	1	ATPA_HAEIN

ALIGNMENTS

RESULT 1

```

PAL_KLEPN
ID PAL_KLEPN STANDARD; PRT; 286 AA.
AC P37446;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DE PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT
DE PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-
DE ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).
GN PLDA.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94131966; PubMed=8300539;
RA Brok R.G.P.M., Brinkman E., van Bostel R., Bekkers A.C.A.P.,
RA Verheij H.M., Tommassen J.;
RT "Molecular characterization of enterobacterial plda genes encoding
RT outer membrane phospholipase A.";
RL J. Bacteriol. 176:861-870(1994).
CC -1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE
CC A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.
CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES
CC LOCATED THERE.
```

```

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X76901; CAA54223.1; -
DR PIR; B36971; B36971.
DR PIR; S40129; S40129.
DR InterPro; IPR003187; PLAL.
DR Pfam; PF02253; PLAL; 1.
KW Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium.
FT SIGNAL 1 20 BY SIMILARITY.
FT CHAIN 21 286 PHOSPHOLIPASE A1.
FT ACT_SITE 161 161 BY SIMILARITY.
SQ SEQUENCE 286 AA; 32544 MW; 3E39F863085108A3 CRC64;
-----
```

Query Match 1.88; Score 8; DB 1; Length 286;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 431 GVGMLND 438
 |||||
 Db 277 GVGMLND 284

RESULT 2

VLTF_BPT5 STANDARD; PRT; 1396 AA.
 AC P13390; 048502;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE L-SHAPED TAIL FIBER PROTEIN (LTF PROTEIN).
 GN LTF.
 OS Bacteriophage T5.
 OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae.
 OC NCBI_TaxID=10726;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95309401; PubMed=7789514;
 RA Kaliman A.V.; Kulshin V.E.; Shlyapnikov M.G.; Ksenzenko V.N.,
 RA Kryukov V.M.;
 RT "The nucleotide sequence of the bacteriophage T5 ltf gene."
 RL FEBS Lett. 366:46-48(1995).
 [2]
 RP SEQUENCE FROM N.A.
 RA Kaliman A.V.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 [3]
 RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
 RX MEDLINE=88289370; PubMed=3267228;
 RA Kaliman A.V.; Kryukov V.M.; Bayev A.A.;
 RT "The nucleotide sequence of bacteriophage T5 DNA at the region
 RT between early and late genes."
 RL Nucleic Acids Res. 16:6230-6230(1988).
 CC -1- FUNCTION: NONESSENTIAL PROTEIN THAT MEDIATES BINDING TO THE
 CC POLYMANNOSE O ANTIGEN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X69460; CAA49220.1; -;
 DR EMBL; AJ001191; CAA04591.1; -;
 DR PIR; S01982; S01982.
 KW Late protein.
 FT CONFLICT 986 986 V -> A (IN REF. 2).
 SQ SEQUENCE 1396 AA; 147989 MW; 18CD2192F65FFFC1 CRC64;

Query Match 1.8%; Score 8; DB 1; Length 1396;
 Best Local Similarity 100.0%; Pred. No. 6.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 54 QSATQSAS 61
 |||||
 Db 87 QSATQSAS 94

RESULT 3

PRL_ICTPU STANDARD; PRT; 186 AA.
 AC P51904;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE PROLACTIN (PRL).
 OS Ictalurus punctatus (Channel catfish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Siluriformes; Ictaluridae; Ictalurus.
 OC NCBI_TaxID=7998;
 [1]
 RP SEQUENCE.

RP TISSUE=Pituitary;
 RX MEDLINE=93364578; PubMed=1308206;
 RA Watanabe K.; Igarashi A.; Noso T.; Chen T.T.; Dunham R.A.,
 RA Kawauchi H.;
 RT "Chemical identification of catfish growth hormone and prolactin."
 RL Mol. Mar. Biol. Biotechnol. 1:239-249(1992).
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
 DR HSSP; Q28632; 1AN3
 DR InterPro; IPR001400; SOMATOTROPIN.
 DR Pfam; PF00103; Hormone; 1.
 DR PRINTS; PR00836; SOMATOTROPIN.
 DR PROSITE; PS00266; SOMATOTROPIN_1; 1.
 DR PROSITE; PS00338; SOMATOTROPIN_2; 1.
 KW Hormone; Pituitary.
 FT DISULFID 45 159 BY SIMILARITY.
 FT DISULFID 176 186 BY SIMILARITY.
 SQ SEQUENCE 186 AA; 20606 MW; 7BCCD89718F44E74 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 186;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 41 NDLGQDN 47
 |||||
 Db 140 NDLGQDN 146

RESULT 4

TBA_PLAYO STANDARD; PRT; 197 AA.
 ID TBA_PLAYO
 AC P12543;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE TUBULIN ALPHA CHAIN (FRAGMENT).
 OS Plasmodium berghei yoelii.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OC NCBI_TaxID=5862;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89014607; PubMed=2459618;
 RA Akella R.; Arasu P.; Vaidya A.B.;
 RT "Molecular clones of alpha-tubulin genes of Plasmodium yoelii reveal
 RT an unusual feature of the carboxy terminus."
 RL Mol. Biochem. Parasitol. 30:165-174(1988).
 CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
 CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
 CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
 CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M29816; AAA29779.1; -;
 DR PIR; A54506; A54506.
 DR InterPro; IPR000217; Tubulin.
 DR InterPro; IPR003008; Tubulin_Ftsz.
 DR Pfam; PF00091; tubulin; 1.
 DR PROSITE; PS00227; TUBULIN; PARTIAL.


```
KW Microtubules; GTP-binding.
FT NON_TER 1
SQ SEQUENCE 197 AA; 22111 MW; 4DAAF199CC6CD319 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 197;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 IKTKRSI 120
DB 82 IKTKRSI 88

RESULT 5
TBA_OCTVU STANDARD; PRT; 240 AA.
AC P24635;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE TUBULIN ALPHA CHAIN (FRAGMENT).
OS Octopus vulgaris (Octopus).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;
OC Incirrata; Octopodidae; Octopus.
OX NCBI_TaxID=6645;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lens;
RA Zinov'Eva R.D., Aleinikova K.S., Tomarev S.I.;
RT "Isolation and structural characterization of cDNAs coding for alpha-
RT tubulin of the octopus eye lens.";
RL Dokl. Akad. Nauk SSSR 302:462-467(1988).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X15845; CAA33844.1; -
DR PIR; A61544; A61544.
DR InterPro; IPR000217; Tubulin.
DR InterPro; IPR003008; Tubulin_Ftsz.
DR Pfam; PF00091; tubulin; 1.
DR PROSITE; PS00227; TUBULIN; PARTIAL.
KW Microtubules; GTP-binding.
FT NON_TER 1
FT SITE 240 240 INVOLVED IN POLYMERIZATION.
SQ SEQUENCE 240 AA; 26961 MW; 13BB3A1F740F2416 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 240;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 IKTKRSI 120
DB 124 IKTKRSI 130

RESULT 6
Y304_MYCPN STANDARD; PRT; 270 AA.
ID Y304_MYCPN
AC P75355;
DT 01-NOV-1997 (Rel. 35, Created)
```

```
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE HYPOPHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN MG304 HOMOLOG
DE (A05_ORF270L).
GN MPN433 OR MP408.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelmreich R., Hilbert H., Plagens H., Pirkil E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000040; AAB96056.1; -
DR InterPro; IPR003593; AAA.
DR InterPro; IPR003439; ABC_transportr.
DR InterPro; IPR001687; ATP_GTP_A.
DR Pfam; PF00005; ABC_tran; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW Hypothetical protein; ATP-binding; Transport; Complete proteome.
FT NP_BIND 36 43 ATP (POTENTIAL).
SQ SEQUENCE 270 AA; 30770 MW; 6513640B4BC051B4 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 270;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 KQFTPLS 174
DB 125 KQFTPLS 131

RESULT 7
ILVE_THEME STANDARD; PRT; 273 AA.
ID ILVE_THEME
AC P74921;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PROBABLE BRANCHED-CHAIN AMINO ACID AMINOTRANSFERASE (EC 2.6.1.42)
DE (BCAT).
GN ILVE OR TM0831.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Steward A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
```

RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 RL genome sequence of *Thermotoga maritima*.";
 RN Nature 399:323-329(1999).
 RP SEQUENCE OF 68-273 FROM N.A.
 RC STRAIN-MSB8 / DSM 3109;
 RX MEDLINE=97017137; PubMed=8863738;
 RA Guipaud O., Labedan B., Forterre P.;
 RT "A gyrB-like gene from the hyperthermophilic bacterium *Thermotoga*
 RL *maritima*.";
 RL Gene 174:121-128(1996).
 CC -1- CATALYTIC ACTIVITY: L-LEUCINE + 2-OXOGlutARATE = 4-METHYL-2-
 CC OXOPENTANOATE + L-GLUTAMATE (ALSO ACTS ON L-ISOLEUCINE AND
 CC L-VALINE).
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC -1- PATHWAY: VALINE AND ISOLEUCINE BIOSYNTHESIS.
 CC -1- SIMILARITY: BELONGS TO CLASS-IV OF PYRIDOXAL-PHOSPHATE-DEPENDENT
 CC AMINOTRANSFERASES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE001750; AAC35913.1; -
 DR EMBL; AU9692; AAC44497.1; -
 DR HSPSP; P00510; IAG3.
 DR TIGR; TM0831; -
 DR InterPro; IPR001544; AminoTran_4.
 DR Pfam; PF01063; aminotran_4; 1.
 DR ProDom; PD001961; AminoTran_4; 1.
 DR PROSITE; PS00770; AA_TRANSFASER_CLASS_4; 1.
 DR Transfaser; AminoTransferase; Branched-chain amino acid biosynthesis;
 KW Pyridoxal phosphate; Complete proteome.
 FT CONFLICT 77 A -> R (IN REF. 2).
 SQ SEQUENCE 273 AA; 31158 MW; 2162B705612E90E3 CRC64;
 Query Match 1.6%; Score 7; DB 1; Length 273;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 172 PLSLSFD 178
 Db 61 PLSLSFD 67
 |||||
 RESULT 8
 ID MIAA_PSEPU STANDARD; PRT; 322 AA.
 AC 030762;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE TRNA DELTA(2)-ISOPENTENYLPROPHOSPHATE TRANSFERASE (EC 2.5.1.8) (TPP
 DE TRANSFERASE) (ISOPENTENYL-DIPHOSPHATE:TRNA ISOPENTENYLTTRANSFERASE)
 DE (IPTASE) (IPT).
 GN MIAA.
 OS Pseudomonas putida.
 CC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 CC Pseudomonas.
 CC NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M;
 RA Olekhnovich I.N., Gussin G.N.;
 RT "Attenuation of the pseudomonas putida trpE and trpGDC genes.";
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CATALYZES THE FIRST STEP IN THE BIOSYNTHESIS OF
 CC 2-METHYLTHIO-N6-(DELTA(2)-ISOPENTENYL)-ADENOSINE (MS[2][6]A)

CC ADJACENT TO THE ANTICODON OF SEVERAL TRNA SPECIES (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ISOPENTENYL DIPHOSPHATE + TRNA -
 CC PYROPHOSPHATE + TRNA CONTAINING 6-ISOPENTENYLADEMOSINE.
 CC -1- SIMILARITY: BELONGS TO THE IPP TRANSFERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF016312; AAB69443.1; -
 DR InterPro; IPR002627; IPPT.
 DR Pfam; PF01715; IPPT; 1.
 DR ProDom; PD004674; IPPT; 1.
 KW Transferase; Nucleotidyltransferase; trna processing; ATP-binding.
 FT NP_BIND 12 19 ATP (POTENTIAL).
 SQ SEQUENCE 322 AA; 35488 MW; DIAA81A2D7B32F6A CRC64;
 Query Match 1.6%; Score 7; DB 1; Length 322;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 33 FVDEVRS 39
 Db 231 FVDEVRS 237
 |||||
 RESULT 9
 ID SPI_BACBR STANDARD; PRT; 326 AA.
 AC P43131;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE 01-NOV-1997 (Rel. 35, Last annotation update)
 DE PROTEASE INHIBITOR PRECURSOR (BBRPI).
 OS Bacillus brevis.
 CC Bacteria; Firmicutes; Bacillus/Clostridium group;
 CC Bacillus/Staphylococcus group; Brevibacillus.
 CC NCBI_TaxID=1393;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-32; 104-112 AND 122-136.
 RC STRAIN=HPD31;
 RX MEDLINE=92304060; PubMed=1610177;
 RA Shiga Y., Hasegawa K., Tsuboi A., Yamagata H., Uda S.;
 RT "Characterization of an extracellular protease inhibitor of *Bacillus*
 RL *brevis* HPD31 and nucleotide sequence of the corresponding gene.";
 RL Appl. Environ. Microbiol. 58:525-531(1992).
 CC -1- FUNCTION: SHOWS INHIBITORY ACTIVITY TOWARDS SERINE PROTEASES, SUCH
 CC AS TRYPSIN, CHYMOTRYPSIN, AND SUBTILISIN. MAY FORM A TRYPSIN-
 CC INHIBITOR COMPLEX IN A MOLAR RATIO OF 1:1. IT IS HEAT RESISTANT AT
 CC NEUTRAL AND ACIDIC PHs.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- PTM: MAY UNDERGO ACTIVATION AFTER SECRETION. IT IS PRODUCED
 CC EXTRACELLULARLY IN MULTIPLE FORMS HAVING AT LEAST THREE DIFFERENT
 CC MOLECULAR WEIGHTS (BBRPI-A, -B, AND -C).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D10696; BAA01538.1; -
 KW Serine protease inhibitor; Protease inhibitor; Signal; Repeat.
 FT SIGNAL 1 24
 FT CHAIN 25 326 SERINE PROTEASE INHIBITOR.
 FT CHAIN 104 326 SERINE PROTEASE INHIBITOR-C.

```

FT CHAIN 104 ? SERINE PROTEASE INHIBITOR-B (POTENTIAL).
FT CHAIN 122 326 SERINE PROTEASE INHIBITOR-A.
FT DOMAIN 177 304 CONTAINS TWO APPROXIMATE REPEATS.
FT REPEAT 177 208 1.
FT REPEAT 272 304 2.
SQ SEQUENCE 326 AA; 35100 MW; 1C0456ABFA912E77 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 326;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 PVAFVDE 36
Db 119 PVAFVDE 125

RESULT 10
ID FLAD_METJA STANDARD; PRT; 342 AA.
AC Q58305;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PUTATIVE FLAGELLA-RELATED PROTEIN D.
GN FLAD OR MJ0895.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kierulff A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67521; AAB98748.1; -
DR TIGR; MJ0755; -
KW Hypothetical protein; Signal; Complete proteome.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 342 HYPOTHETICAL PROTEIN MJ0755.
SQ SEQUENCE 342 AA; 38778 MW; BD25220A7EAD85C CRC64;

Query Match 1.6%; Score 7; DB 1; Length 342;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 YTTALEN 82
Db 90 YTTALEN 96

RESULT 12
NU2M_CAPHI STANDARD; PRT; 347 AA.
ID NU2M_CAPHI
AC Q36346;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 2 (EC 1.6.5.3).
GN MTND2 OR ND2.
OS Capra hircus (Goat).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RC Doxv P., Mann W., Hecht W.;
RT TISSUE=Liver;
RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE -> NAD(+) + UBIQUINOL.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL

```

```

CC CC INNER MEMBRANE.
CC CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 2 FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
DR EMBL; X72965; CAA51468.1; -.
DR InterPro; IPR003917; NADHUB_oxdrdctse2.
DR InterPro; IPR001750; Oxidored_q1.
DR Pfam; PF00361; oxidored_q1; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
SQ SEQUENCE 347 AA; 39136 MW; 6129CFF90980475E CRC64;

Query Match 1.6%; Score 7; DB 1; Length 347;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 LTLTSL 14
Db 155 LTLTSL 161

RESULT 13
NU2M_HUMAN STANDARD; PRT; 347 AA.
AC P03891; Q34769; Q9TGI0; Q9TGI1; Q9TGI2; Q9TGI3; Q9TGI4;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 2 (EC 1.6.5.3).
GN MTND2 OR ND2.
OS Homo sapiens (Human).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81173052; PubMed=7219534;
RA Anderson S., Bankier A.T., Barrell B.G., de Bruijn M.H.L.,
RA Coulson A.R., Drouin J., Eperon I.C., Nierlich D.P., Roe B.A.,
RA Sanger F., Schreier P.H., Smith A.J.H., Staden R., Young I.G.;
RT "Sequence and organization of the human mitochondrial genome.";
RL Nature 290:457-465(1981).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=81170577; PubMed=6260957;
RA Sanger F., Coulson A.R., Barrell B.G., Smith A.J.H., Roe B.A.;
RT "Cloning in single-stranded bacteriophage as an aid to rapid DNA
RT sequencing.";
RL J. Mol. Biol. 143:161-178(1980).
RN [3]
RP SEQUENCE FROM N.A. AND VARIANTS.
RX MEDLINE=98133898; PubMed=9475751;
RA Wise C.A., Stram M., Eastel S.;
RT "Departure from neutrality at the mitochondrial NADH dehydrogenase
RT subunit 2 gene in humans, but not in chimpanzees.";
RL Genetics 148:409-421(1998).
RN [4]
RP SEQUENCE FROM N.A. AND VARIANTS ILE-43; LEU-325 AND THR-331.
RC TISSUE=Placenta;
RX MEDLINE=95132634; PubMed=7530363;
RA Horai S., Hayasaka K., Kondo R., Tsugane K., Takahata N.;
RT "Recent African origin of modern humans revealed by complete sequences
RT of hominoid mitochondrial DNAs.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:532-536(1995).
RN [5]

```

```

RP RP IDENTIFICATION OF PROTEIN.
RX RX MEDLINE=85188293; PubMed=3921850;
RA Chomyn A., Mariottini P., Cleeter M.W.J., Ragan C.I., Matsuno-Yagi A.,
RA Hatefi Y., Doolittle R.F., Attardi G.;
RT "Six unidentified reading frames of human mitochondrial DNA encode
RT components of the respiratory-chain NADH dehydrogenase.";
RL Nature 314:592-597(1985).
RN [6]
RP VARIANT LHON ASP-150.
RX MEDLINE=91144615; PubMed=19000003;
RA Johns D.R., Berman J.;
RT "Alternative, simultaneous complex I mitochondrial DNA mutations in
RT Leber's hereditary optic neuropathy.";
RL Biochem. Biophys. Res. Commun. 174:1324-1330(1991).
RN [7]
RP VARIANTS L-42; R-63; A-119; P-148; S-150; T-159 AND A-185.
RX MEDLINE=92098084; PubMed=1757091;
RA Marzuki S., Noer A.S., Lertrit P., Thyagarajan D., Kapsa R.,
RA Utthanaphol P., Byrne E.;
RT "Normal variants of human mitochondrial DNA and translation products:
RT the building of a reference data base.";
RL Hum. Genet. 88:139-145(1991).
RN [8]
RP VARIANT LHON SER-259.
RX MEDLINE=92120513; PubMed=1732158;
RA Brown M.D., Voljavec A.S., Lott M.T., Torroni A., Yang C.C.,
RA Wallace D.C.;
RT "Mitochondrial DNA complex I and III mutations associated with
RT Leber's hereditary optic neuropathy.";
RL Genetics 130:163-173(1992).
RN [9]
RP VARIANT AD SER-331.
RX MEDLINE=92118019; PubMed=1370613;
RA Lin F.-H., Lin R., Wisniewski H.M., Hwang Y.-W., Grundke-Iqbal I.,
RA Healy-Louie G., Iqbal K.;
RT "Detection of point mutations in codon 331 of mitochondrial NADH
RT dehydrogenase subunit 2 in Alzheimer's brains.";
RL Biochem. Biophys. Res. Commun. 182:238-246(1992).
RN [10]
RP VARIANT THR-57.
RX MEDLINE=98127994; PubMed=9461455;
RA Rieder M.J., Taylor S.L., Tobe V.O., Nickerson D.A.;
RT "Automating the identification of DNA variations using quality-based
RT fluorescence re-sequencing: analysis of the human mitochondrial
RT genome.";
RL Nucleic Acids Res. 26:967-973(1998).
CC CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE -> NAD(+) + UBIQUINOL.
CC CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC CC INNER MEMBRANE.
CC CC -1- DISEASE: DEFECTS IN MTND2 ARE ONE OF THE CAUSES OF LEBER'S
CC CC HEREDITARY OPTIC NEUROPATHY (LHON), A MATERNALLY INHERITED DISEASE
CC CC RESULTING IN ACUTE BILATERAL BLINDNESS DUE TO RETINAL DEGENERATION
CC CC PREDOMINANTLY IN YOUNG MEN. CARDIAC CONDUCTION DEFECTS AND
CC CC NEUROLOGICAL DEFECTS HAVE ALSO BEEN DESCRIBED. RESULTING IN OPTIC
CC CC NERVE DEGENERATION AND CARDIAC DYSRHYTHMIA.
CC CC -1- DISEASE: DEFECTS IN MTND2 COULD BE ASSOCIATED WITH ALZHEIMER'S
CC CC DISEASE (AD).
CC CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 2 FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
DR EMBL; J01415; AAB58944.1; -.
DR EMBL; V00662; CAA24027.1; -.
DR EMBL; M10546; AAA65502.1; ALT_INIT.
DR EMBL; D38112; BAA07291.1; -.
DR EMBL; AF014882; AAC25441.1; -.
DR EMBL; AF014884; AAC25443.1; -.

```


P92691;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 2 (EC 1.6.5.3).
 GN MTND2 OR ND2 OR NADH2.
 OS Pongo pygmaeus abelii (Sumatran orangutan).
 OG Mitochondrion.
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pongo.
 OX NCBI_TaxID=9601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YN93-312;
 RX MEDLINE=97032590; PubMed=8875856;
 RA Xu X., Arnason U.;
 RT "The mitochondrial DNA molecule of Sumatran orangutan and a molecular
 RT proposal for two (Bornean and Sumatran) species of orangutan.";
 RL J. Mol. Evol. 43:431-437(1996).
 CC -|- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
 CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 CC INNER MEMBRANE.
 CC -|- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 2 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X97707; CAA66284.1; .
 DR InterPro: IPR003917; NADhub_oxdrdctse2.
 DR InterPro: IPR001750; Oxidored_q1.
 DR Pfam: PF00361; oxidored_q1; 1.
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
 SQ SEQUENCE 347 AA; 38619 MW; 53885266466A2B93 CRC64;

Query Match 1.68; Score 7; DB 1; Length 347;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 8 LTLSILS 14
 |||||
 Db 155 LTLSILS 161

Search completed: May 9, 2002, 13:05:51
 Job time: 606 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:55:05 ; Search time 77.39 Seconds
(without alignments)
835.410 Million cell updates/sec

Title: US-09-787-083-2
Perfect score: 442
Sequence: 1 MKVSILSTLTLSCFAILA.....YNHEATSGVGLMLNDWML 442

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 473505 seqs, 146272329 residues

Word size : 0

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL_17:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organella:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	2.0	382	2 Q9K0U7	Q9K0U7 neisseria m
2	9	2.0	409	2 Q9J721	Q9J721 neisseria m
3	9	2.0	862	2 Q9C162	Q9C162 pasteurella
4	9	2.0	1432	10 Q9AX85	Q9AX85 oryza sativ
5	8	1.8	299	10 Q9W268	Q9W268 arabidopsis
6	8	1.8	306	2 Q9CL22	Q9CL22 pasteurella
7	7	1.6	55	12 Q67608	Q67608 tomato gold
8	7	1.6	72	12 Q9IZS5	Q9IZS5 potato viru
9	7	1.6	72	12 Q9IZS4	Q9IZS4 potato viru
10	7	1.6	95	12 Q9IZS3	Q9IZS3 potato viru
11	7	1.6	96	12 Q9IZS8	Q9IZS8 potato viru
12	7	1.6	100	3 Q12148	Q12148 saccharomyc
13	7	1.6	101	12 Q9IZS6	Q9IZS6 potato viru
14	7	1.6	122	2 Q9RU16	Q9RU16 deinococcus
15	7	1.6	124	12 Q9IZS7	Q9IZS7 potato viru
16	7	1.6	131	6 Q9GMD4	Q9GMD4 macaca fusc
17	7	1.6	138	12 Q9IZQ7	Q9IZQ7 potato viru
18	7	1.6	147	5 Q36035	Q36035 hexamita in
19	7	1.6	158	2 Q9Z9C1	Q9Z9C1 chlamydia p

20	7	1.6	169	2	Q9L2A7	Q9L2A7 streptomyce
21	7	1.6	186	10	Q9LRX9	Q9LRX9 arabidopsis
22	7	1.6	189	13	Q9DFT2	Q9DFT2 notothenia
23	7	1.6	198	4	Q9BUX9	Q9BUX9 homo sapien
24	7	1.6	212	13	Q9I819	Q9I819 ictalurus p
25	7	1.6	214	5	Q9VTR5	Q9VTR5 drosophila
26	7	1.6	216	2	Q9X6H6	Q9X6H6 streptococc
27	7	1.6	218	2	Q9RF14	Q9RF14 streptococc
28	7	1.6	219	5	O17594	O17594 caenorhabdi
29	7	1.6	220	2	O9XCK1	O9XCK1 streptococc
30	7	1.6	220	5	O36036	O36036 spironucleu
31	7	1.6	221	12	Q85002	Q85002 alifalra mos
32	7	1.6	227	5	Q9BMF4	Q9BMF4 glossina mo
33	7	1.6	233	6	Q9MZB0	Q9MZB0 ovis aries
34	7	1.6	236	13	Q9DFS8	Q9DFS8 notothenia
35	7	1.6	264	12	Q85260	Q85260 potato viru
36	7	1.6	265	9	Q38155	Q38155 bacterioph
37	7	1.6	267	2	Q9K732	Q9K732 bacillus ha
38	7	1.6	267	9	Q99363	Q99363 bacterioph
39	7	1.6	267	12	Q85261	Q85261 potato viru
40	7	1.6	267	12	Q98628	Q98628 potato viru
41	7	1.6	267	12	Q86842	Q86842 potato viru
42	7	1.6	267	12	Q99026	Q99026 potato viru
43	7	1.6	267	12	Q05476	Q05476 potato viru
44	7	1.6	267	12	O09711	O09711 potato viru
45	7	1.6	267	12	Q9QNM9	Q9QNM9 potato viru

ALIGNMENTS

RESULT 1
Q9K0U7 PRELIMINARY; PRT; 382 AA.
AC Q9K0U7;
DT 01-OCT-2000 (TREMREL. 15, Created)
DT 01-OCT-2000 (TREMREL. 15, Last sequence update)
DT 01-JUN-2001 (TREMREL. 17, Last annotation update)
DE PHOSPHOLIPASE A1, PUTATIVE.
GN NMB0464;
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J., Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A., Mason T., Ciecko A., Parksey D.S., Blair E., Citterone H., Clark E.B., Cotton M.D., Uterback T.R., Khouri H., Qin H., Vamathevan J., Gill J., Scarlatti V., Masignani V., Pizzo M., Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain MC58";
RL Science 287:1809-1815(2000).
DR EMBL: AE002403; AAF40901.1; -.
DR TIGR: NMB0464; -.
DR InterPro: IPR003187; PLAL.
DR Pfam: PF02253; PLAL; 1.
KW Complete proteome.
SQ SEQUENCE 382 AA; 42714 MW; B468A802F062E836 CRC64;

Query Match 2.0%; Score 9; DB 2; Length 382;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 279 PEIFLTPQPV 287

Db 221 PEIFLTPQPV 229

```

RESULT 2
Q9J721 ID Q9J721 PRELIMINARY; PRT; 409 AA.
AC Q9J721
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE PUTATIVE PHOPHOLIPASE.
GN NMA2021.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis 22491.";
RL Nature 404:502-506(2000).
DR EMBL: AL162757; CAB85240.1; -.
DR InterPro; IPR003187; PLAL.
DR Pfam; PF02253; PLAL; 1.
KW Complete proteome.
SQ SEQUENCE 409 AA; 45862 MW; CD6585B064D01A41 CRC64;

Query Match 2.0%; Score 9; DB 2; Length 409;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEFLTPQV 287
DB 248 PEFLTPQV 256

RESULT 3
Q9CL62 ID Q9CL62 PRELIMINARY; PRT; 862 AA.
AC Q9CL62
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CVGSY.
GN CVGSY OR PM1380.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -!- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
CC REGULATORY COMPONENTS OF SENSORY TRANSDUCTION SYSTEMS.
CC -!- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
CC KINASES.
DR EMBL: AE006176; AAK03464.1; -.
DR InterPro; IPR000410; Bctrl_sensor.
DR InterPro; IPR000658; DUF5.
DR InterPro; IPR003660; HAMP.
DR InterPro; IPR003594; HATPase_c.
DR InterPro; IPR003661; His_kinA.

InterPro; IPR001789; Response_reg.
DR Pfam; PF00672; DUF5; 1.
DR Pfam; PF00072; response_reg; 1.
DR Pfam; PF00512; signal; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; HSKA; 1.
DR SMART; SM00448; REC; 1.
KW Complete proteome; Kinase; Phosphorylation; Sensory transduction;
KW Transferase.
SQ SEQUENCE 862 AA; 97846 MW; AB6814A3B63626C0 CRC64;

Query Match 2.0%; Score 9; DB 2; Length 862;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LSTLTLSIL 13
DB 293 LSTLTLSIL 301

RESULT 4
Q9AX85 ID Q9AX85 PRELIMINARY; PRT; 1432 AA.
AC Q9AX85
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE PUTATIVE ABC TRANSPORTER PROTEIN.
GN P0410E03.7.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0410E03.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002844; BAB21276.1; -.
SQ SEQUENCE 1432 AA; 160234 MW; 04E8CF2CAA08D993 CRC64;

Query Match 2.0%; Score 9; DB 10; Length 1432;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 FTPLSLSF 178
DB 830 FTPLSLSF 838

RESULT 5
Q9M268 ID Q9M268 PRELIMINARY; PRT; 299 AA.
AC Q9M268
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE RAV-LIKE PROTEIN.
GN F21F14.140.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
```


RA Choisine N., Robert C., Brottier P., Wincker P., Wincker P., Cattolico L.,
RA Artyguenave F., Saurin W., Weissenbach J., Mewes H.W., Lemcke K.,
RA Mayer K.F.X., Quetier F., Salanoubat M.,
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; ALI38642; CAB71904.1; -
DR InterPro; IPR003340; B3.
DR Pfam; PF02362; B3; 1.
SQ SEQUENCE 299 AA; 34270 MW; 2AB841F179DD174B CRC64;

Query Match 1.8%; Score 8; DB 10; Length 299;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 174 SLSFDLDR 181
| | | | |
Db 292 SLSFDLDR 299

RESULT 6
Q9CL22 PRELIMINARY; PRT; 306 AA.
ID Q9CL22
AC Q9CL22
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL PROTEIN PM1426.
GN PM1426.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL; AE006179; AA03510.1; -
DR InterPro; IPR003187; PLA1.
DR Pfam; PF02253; PLA1; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 306 AA; 35580 MW; EAF3DE8C1C22B26E CRC64;

Query Match 1.8%; Score 8; DB 2; Length 306;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 314 KLSRSNR 321
| | | | |
Db 184 KLSRSNR 191

RESULT 7
Q67608 PRELIMINARY; PRT; 55 AA.
ID Q67608
AC Q67608;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SUBGENOMIC DNA DERIVED FROM DNA B CCDS - COVALENTLY CLOSED CIRCULAR
DE DOUBLE-STRANDED MOLECULE.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10831;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=87040767; PubMed=3022243;
RA MacDowell S.W., Coutts R.H.A., Buck K.W.;
RT "Molecular characterisation of subgenomic single-stranded and double-
RT stranded DNA forms isolated from plants infected with tomato golden
RT mosaic virus.";
RL Nucleic Acids Res. 14:7967-7984(1986).
DR EMBL; X04485; CAA28171.1; -
DR InterPro; IPR00211; Gemini_BL.
DR Pfam; PF00845; Gemini_BL1; 1.
SQ SEQUENCE 55 AA; 6281 MW; 41963570D739EC96 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 55;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 312 SAKLSRS 318
| | | | |
Db 43 SAKLSRS 49

RESULT 8
Q9IZS5 PRELIMINARY; PRT; 72 AA.
ID Q9IZS5
AC Q9IZS5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y strain NTN (PVY(NTN)).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=122280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUBER NECROSING (NTN);
RA Cerovska N., Filigarova M., Moravec T., Petrzik K.;
RT "Differences in nucleotide and amino acid sequences of N-terminal
RT parts of coat proteins among isolates of potato virus Y-NTN strain.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF28633; AAF67874.1; -
FT NON_TER 1
FT NON_TER 72
SQ SEQUENCE 72 AA; 7721 MW; 2C5287E15624CBDD CRC64;

Query Match 1.6%; Score 7; DB 12; Length 72;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 156 TKDKAQ 162
| | | | |
Db 11 TKDKAQ 17

RESULT 9
Q9IZS4 PRELIMINARY; PRT; 72 AA.
ID Q9IZS4
AC Q9IZS4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y strain NTN (PVY(NTN)).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=122280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUBER NECROSING (NTN);
RA Cerovska N., Filigarova M., Moravec T., Petrzik K.;

RT "Differences in nucleotide and amino acid sequences of N-terminal
RT parts of coat proteins among isolates of potato virus Y-NTN strain.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF228634; AAF67875.1; -.
FT NON_TER 1
SQ SEQUENCE 72 AA; 7692 MW; 4F129B870523B482 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 72;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKQ 162
Db 11 TKKDAKQ 17
|||||

RESULT 10
Q91ZS3 PRELIMINARY; PRT; 95 AA.
AC Q91ZS3
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y strain NTN (PVY(NTN)).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=122280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TUBER NECROSING (NTN);
RA Cerovska N., Filligatova M., Moravec T., Petrzik K.;
RT "Differences in nucleotide and amino acid sequences of N-terminal
RT parts of coat proteins among isolates of potato virus Y-NTN strain.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF228635; AAF67876.1; -.
FT NON_TER 1
SQ SEQUENCE 95 AA; 10446 MW; 28C06891777C9928 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 95;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKQ 162
Db 11 TKKDAKQ 17
|||||

RESULT 11
Q91ZS8 PRELIMINARY; PRT; 96 AA.
AC Q91ZS8
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y (strain N) (PVY).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=12219;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NECROSING (N);
RA Cerovska N., Filligatova M., Moravec T., Petrzik K.;
RT "Differences in nucleotide and amino acid sequences of N-terminal
RT parts of coat proteins among isolates of potato virus Y-NTN strain.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF228630; AAF67871.1; -.
FT NON_TER 1
SQ SEQUENCE 96 AA; 10559 MW; 89C3A5675B9784CA CRC64;

Query Match 1.6%; Score 7; DB 12; Length 96;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKQ 162
Db 11 TKKDAKQ 17
|||||

RESULT 12
Q12148 PRELIMINARY; PRT; 100 AA.
AC Q12148
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE ORF YDL163W.
GN YDL163W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Pohl T.M.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MIPS;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-ALPHAS288C;
RA Pohl T.M.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z74212; CAA98738.1; -.
DR EMBL: Z67750; CAA91583.1; -.
DR SGD; S0002322; YDL163W.
KW Hypothetical protein.
SQ SEQUENCE 100 AA; 10751 MW; 952FB92C852AA710 CRC64;

Query Match 1.6%; Score 7; DB 3; Length 100;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 LTLSTLS 14
Db 26 LTLSTLS 32
|||||

RESULT 13
Q91ZS6 PRELIMINARY; PRT; 101 AA.
AC Q91ZS6
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y strain NTN (PVY(NTN)).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=122280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TUBER NECROSING (NTN);

```
RA Cerovska N., Filigarova M., Moravec T., Petrzik K.;
RT "Differences in nucleotide and amino acid sequences of N-terminal
RL parts of coat proteins among isolates of potato virus Y-NTN strain.";
DR EMBL: AF228632; AAF67873.1; -
FT NON_TER 101
FT NON_TER 101
SQ SEQUENCE 101 AA; 11112 MW; 855BF51C3D5FE9BD CRC64;

Query Match 1.68; Score 7; DB 12; Length 101;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKDKAQ 162
DB 11 TKDKAQ 17

RESULT 14
Q9RU16 PRELIMINARY; PRT; 122 AA.
AC Q9RU16;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL 13.7 KDA PROTEIN.
GN DR1580.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RC STRAIN=RL;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL: AE002002; AAF11150.1; -
DR TIGR: DR1580; -
DR InterPro: IPR000325; Glyoxalase_1.
DR Pfam: PF00903; Glyoxalase; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 122 AA; 13683 MW; 7BA836DCAD786E51 CRC64;

Query Match 1.68; Score 7; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 EHEPELY 76
DB 115 EHEPELY 121

RESULT 15
Q9IZS7 PRELIMINARY; PRT; 124 AA.
AC Q9IZS7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y strain NTN (PVY(NTN)).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
```

```
OC Potyvirus.
OX NCBI_TaxID=122280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUBER NECROSING (NTN);
RA Cerovska N., Filigarova M., Moravec T., Petrzik K.;
RT "Differences in nucleotide and amino acid sequences of N-terminal
RL parts of coat proteins among isolates of potato virus Y-NTN strain.";
DR EMBL: AF228631; AAF67872.1; -
FT NON_TER 124
FT NON_TER 124
SQ SEQUENCE 124 AA; 13654 MW; DE3618D620BA7080 CRC64;

Query Match 1.68; Score 7; DB 12; Length 124;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKDKAQ 162
DB 11 TKDKAQ 17

Search completed: May 9, 2002, 13:05:17
Job time: 612 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:54:58 ; Search time 72.68 Seconds
(without alignments)
450.473 Million cell updates/sec

Title: US-09-787-083-4

Perfect score: 442

Sequence: 1 MKVSLTLTLILPCFAILA.....YNHEATSGVGLMLNDWMGL 442

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 522463 seqs, 74073290 residues

Word size : 0

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

A_Geneseq_1101.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	442	100.0	442	21	AAV85269 BASB034 amino acid
2	390	88.2	442	21	AAV85270 BASB034 amino acid
3	289	65.4	442	21	AAV85271 BASB034 amino acid
4	224	50.7	442	21	AAV85268 BASB034 amino acid
5	9	2.0	370	21	AAV75156 Neisseria gonorrhoe
6	9	2.0	370	21	AAV75157 Neisseria meningit
7	9	2.0	370	21	AAV75158 Neisseria meningit
8	9	2.0	374	21	AAV70629 Neisseria meningit
9	9	2.0	375	21	AAV70628 Neisseria meningit
10	8	1.8	278	21	AAV11137 Arabidopsis thalia
11	8	1.8	283	21	AAV11136 Arabidopsis thalia

12	1.8	299	21	AAV11135 Arabidopsis thalia
13	1.6	16	15	AAV52127 Mouse light chain
14	7	50	21	AAV79483 Rat alpha tubulin
15	7	64	21	AAV06322 Arabidopsis thalia
16	7	82	21	AAV01488 Human secreted pro
17	7	106	21	AAV07608 Arabidopsis thalia
18	7	106	21	AAV51952 Arabidopsis thalia
19	7	108	21	AAV37811 Arabidopsis thalia
20	7	116	21	AAV87313 Human signal pepti
21	7	136	21	AAV37810 Arabidopsis thalia
22	7	169	20	AAV34661 Chlamydia pneumoni
23	7	181	21	AAV37809 Arabidopsis thalia
24	7	186	21	AAV07607 Arabidopsis thalia
25	7	186	21	AAV51951 Arabidopsis thalia
26	7	188	21	AAV51950 Arabidopsis thalia
27	7	190	21	AAV07606 Arabidopsis thalia
28	7	268	13	AAV21542 Envelope protein o
29	7	302	13	AAV24254 Protease inhibitor
30	7	316	21	AAV39500 Arabidopsis thalia
31	7	324	21	AAV39499 Arabidopsis thalia
32	7	326	13	AAV24356 Protease inhibitor
33	7	365	21	AAV39498 Arabidopsis thalia
34	7	381	20	AAV60241 Human endometrium
35	7	448	21	AAV91960 Human cytoskeleton
36	7	523	21	AAV44285 Arabidopsis thalia
37	7	552	16	AAV67007 Firefly luciferase
38	7	666	20	AAV21871 Amino acid sequenc
39	7	683	20	AAV21869 Amino acid sequenc
40	7	684	20	AAV21873 Amino acid sequenc
41	7	684	20	AAV21875 Amino acid sequenc
42	7	715	20	AAV21860 Amino acid sequenc
43	7	783	20	AAV70586 Human SLAP-130. H
44	7	900	20	AAV96264 Human axin. Homo
45	7	934	17	AAV07637 P. suis leukotoxin

ALIGNMENTS

RESULT 1

AAV85269

ID AAV85269 standard; Protein: 442 AA.

XX AC AAV85269;

XX AC AAV85269;

DT 29-JUN-2000 (first entry)

DE BASB034 amino acid sequence #2.

XX KW Moraxella catarrhalis infection; BASB034; diagnosis; staging;

XX KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;

XX KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;

XX KW hearing loss; antibacterial drug.

XX OS Moraxella catarrhalis.

XX PN WO200015802-A1.

PD 23-MAR-2000.

XX 14-SEP-1999; 99WO-EP06781.

XX 14-SEP-1998; 98GB-0020002.

XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX Ruelle J;

XX WPI; 2000-271440/23.

XX N-PSDB: AAA10701.

XX Novel BASB034 polynucleotides and polypeptides from Moraxella

XX catarrhalis used to prepare vaccines against bacterial infections

PT

XX PS Claim 3; Page 67; 106pp; English.

XX CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from

CC strain Mc2908. The invention relates to BASB034 polypeptides from

CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034

CC polynucleotides and polypeptides may be employed as research reagents and

CC material for the discovery of treatments and diagnostics for diseases,

CC particularly human diseases. They are particularly used to diagnose and

CC treat M. catarrhalis infections. They can be used for diagnosis of

CC disease, staging of disease, or determining response of an infectious

CC organism to drugs. The polynucleotides may be used as a source for

CC hybridization probes, and for screening of genetic mutations, serotype,

CC organism or strain identification, identification of mutations in BASB034

CC sequences, and as components of arrays which are useful for diagnostic

CC and prognostic purposes. The polypeptides can be used to produce

CC antibodies. The polypeptides can also be used in vaccine formulations,

CC and to identify agonists and antagonists. The polypeptides, antibodies,

CC agonists and antagonists (which are bacteriostatic) are used for the

CC treatment and prevention of diseases such as otitis media in infants and

CC children, pneumonia in elderlies, sinusitis, nosocomial infections and

CC invasive diseases, and chronic otitis media with hearing loss. The

CC polypeptides, agonists and antagonists are also used for screening of

CC antibacterial drugs. The BASB034 products of the invention can be used

CC screen for new antibacterial compounds that may target resistant

CC bacteria.

XX CC

XX SQ Sequence 442 AA;

Query Match 100.0%; Score 442; DB 21; Length 442;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVSLTLPLSLPFAILLAQQAQVNPVAFVDEVRSKNDLQDNELLIGVQATQSA 60

Db 1 mkvsltlslplcfailaiqqagavnpvafvdevrskndlgqdnelligvqatqsa 60

QY 61 SYDTANPLDEHEPELYTTALENKTMLINCSALNQDLMRLACYDTLVHGTPAVIKTKRSI 120

Db 61 stdtanpldehepeelyttalenktmlinclsalnqdmrlacydltlvhgtpaviktksri 120

QY 121 RDETIWTIKGKPVVYQETDPIFLMGNEKGMILTKDAKOLEYAAKQFTPLSLSFDD 180

Db 121 ridetiwtikgkpvvyqetdtpiflmgnekgmttkdakqleyaakqftplsfsdld 180

QY 181 RNTPLWSSRRPHNPMYVLPFIEMHGKPNRSPNTPSHEARQFTNEFRAPLKFQVSVKVA 240

Db 181 rntplwssrrphnmpyvlpifmhgkpnrsnptpshearqftnefrapelkfqvsvkva 240

QY 241 AEDLWGTSDSLWFGYTTQSHWQIFNGKNSRPRVRVHDYQPEIFLTQPVYSDLPWDGKVRMI 300

Db 241 aedlwgtstdslwfgytqshwqifngknsrprvrhdyqpeifltqpvysdlpwdgkvrm 300

QY 301 GNGAVHHSNGESAKLSRSNNRAYLMAGMEWKNLTVMNPRIWGRIFKEGSGSQPDNDPILD 360

Db 301 gngavhhsngesaklsrsnnraylmagmewknlvmnpriwgrifkegsgsqpdndpild 360

QY 361 YXGVDVRELYOLENKSNTSGVRYNPRSGKALQLDVYVPLGKISGVFQIFQGYGQSL 420

Db 361 yygvgdvrelyolenksntsgvrynprrsgkalqldvyyvplgkigsvfqi fgygqsl 420

QY 421 IDYNHEATSFVGLMLNDWML 442

Db 421 idynheatsfvglmlndwml 442

RESULT 2

ID AAY85270

XX AAY85270 standard; Protein; 442 AA.

AC AAY85270;

XX

DT 29-JUN-2000 (first entry)

XX BASB034 amino acid sequence #3.

DE

XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;

KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;

KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;

KW hearing loss; antibacterial drug.

XX

OS Moraxella catarrhalis.

XX

XX WO200015802-A1.

PN

XX 23-MAR-2000.

PD

XX 14-SEP-1999; 99WO-EP06781.

PF

XX 14-SEP-1998; 98GB-0020002.

PR

XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

PA

XX Ruelle J;

PI

XX WPI; 2000-271440/23.

XX

DR N-PSDB; AAA10702.

DR

XX

PT Novel BASB034 polynucleotides and polypeptides from Moraxella

XX catarrhalis used to prepare vaccines against bacterial infections

PS Claim 3; Page 68; 106pp; English.

XX

CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from

CC strain Mc2913. The invention relates to BASB034 polypeptides from

CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034

CC polynucleotides and polypeptides may be employed as research reagents and

CC material for the discovery of treatments and diagnostics for diseases,

CC particularly human diseases. They are particularly used to diagnose and

CC treat M. catarrhalis infections. They can be used for diagnosis of

CC disease, staging of disease, or determining response of an infectious

CC organism to drugs. The polynucleotides may be used as a source for

CC hybridization probes, and for screening of genetic mutations, serotype,

CC organism or strain identification, identification of mutations in BASB034

CC sequences, and as components of arrays which are useful for diagnostic

CC and prognostic purposes. The polypeptides can be used to produce

CC antibodies. The polypeptides can also be used in vaccine formulations,

CC and to identify agonists and antagonists. The polypeptides, antibodies,

CC agonists and antagonists (which are bacteriostatic) are used for the

CC treatment and prevention of diseases such as otitis media in infants and

CC children, pneumonia in elderlies, sinusitis, nosocomial infections and

CC invasive diseases, and chronic otitis media with hearing loss. The

CC polypeptides, agonists and antagonists are also used for screening of

CC antibacterial drugs. The BASB034 products of the invention can be used

CC screen for new antibacterial compounds that may target resistant

CC bacteria.

XX CC

XX SQ Sequence 442 AA;

Query Match 88.2%; Score 390; DB 21; Length 442;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 VQSQTASASTDTANPLDEHEPELYTTALENKTMLINCSALNQDLMRLACYDTLVHGTEPA 112

Db 53 vqsqtasastdtanpldehepeelyttalenktmlinclsalnqdmrlacydltlvhgetpa 112

QY 113 VIKTKRSIRLDETIWTIKGKPVVYQETDPIFLMGNEKGMILTKDAKQLEYAAKQFTP 172

Db 113 viktkrsirldetiwtikgkpvvyqetdtpiflmgnekgmttkdakqleyaakqfcp 172

QY 173 LSLSFOLDLRNNTPLWSSRRPHNPMYVLPFIEMHGKPNRSPNTPSHEARQFTNEFRAPLKF 232

Db 173 lslsfoldrntplwssrrphnmpyvlpifmhgkpnrsnptpshearqftnefrapelkf 232

QY 233 QVSVKAAEDLWGTDSLWFGYTQQSHWQIFNGKNSRPRVRHDYQPEIFLTQPVYSDLP 292
|||||
Db 233 qvsvkkaeedlwgtdsdlwfgytqqshwqifngknsrprvrhdypqeilftqpvysdpl 292
|||||
QY 293 WDKVRMIGMGAHVHSGESAKLSRSWNRAYLMAGMEWKNLTVMPRTWGRIFKEGSGSQP 352
|||||
Db 293 wdkvrmi gmga hvhsgesaklsrswnraylmagmewknl tvmprtwgrifkegsgsqp 352
|||||
QY 353 DDNPDI LDYGYGDVRFYOLENKSNTSGTVRYNPRSGKALOLDYVYPLGKISGYFOI 412
|||||
Db 353 ddnpdildygygdvrfyolenksntsgtvrynp rsgkaloldyvyplgk isgyfqi 412
|||||
QY 413 FQYGGSLIDYNHEATSFVGGLMLNDWMGL 442
|||||
Db 413 fqyggslidynheatsfgvglmndwmgl 442
|||||

RESULT 3
AAY85271
ID AAY85271 standard; Protein; 442 AA.
XX
AC AAY85271;
XX
DT 29-JUN-2000 (first entry)
XX
DE BASB034 amino acid sequence #4.
XX
KW Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW hearing loss; antibacterial drug.
XX
OS Moraxella catarrhalis.
XX
PN WO200015802-A1.
XX
PD 23-MAR-2000.
XX
PF 14-SEP-1999; 99WO-EP06781.
XX
PR 14-SEP-1998; 98GB-0020002.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Ruelle J;
XX
DR WPI; 2000-271440/23.
DR N-PSDB; AAA10703.
XX
PT Novel BASB034 polynucleotides and polypeptides from Moraxella
PT catarrhalis used to prepare vaccines against bacterial infections -
XX
PS Claim 3; Page 69; 106pp; English.

This sequence represents a Moraxella catarrhalis BASB034 polypeptide from strain Mc2969. The invention relates to BASB034 polypeptides from M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034 polynucleotides and polypeptides may be employed as research reagents and material for the discovery of treatments and diagnostics for diseases, particularly human diseases. They are particularly used to diagnose and treat M. catarrhalis infections. They can be used for diagnosis of disease, staging of disease, or determining response of an infectious organism to drugs. The polynucleotides may be used as a source for hybridization probes, and for screening of genetic mutations, serotype, organism or strain identification, identification of mutations in BASB034 sequences, and as components of arrays which are useful for diagnostic and prognostic purposes. The polypeptides can be used to produce antibodies. The polypeptides can also be used in vaccine formulations, and to identify agonists and antagonists. The polypeptides, antibodies, agonists and antagonists (which are bacteriostatic) are used for the treatment and prevention of diseases such as otitis media in infants and children, pneumonia in elderlies, sinusitis, nosocomial infections and

CC Invasive diseases, and chronic otitis media with hearing loss. The
CC polypeptides, agonists and antagonists are also used for screening of
CC antibacterial drugs. The BASB034 products of the invention can be used
CC screen for new antibacterial compounds that may target resistant
CC bacteria.
XX
SQ Sequence 442 AA;
Query Match 65.4%; Score 289; DB 21; Length 442;
Best Local Similarity 99.7%; Pred. No. 7.6e-273;
Matches 389; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 53 VQSATQSASTDTANPLDEHEPELYTTALENKTMLINCNSALNQDITMLACYDTLVHGETPA 112
|||||
Db 53 vqsatqsastdtanpldehepe lyttalenktmlincnsalnqdimriacydtlvhgetpa 112
|||||
QY 113 VIKTKRSIRLDETITWOTIKGKQPVVYOETTDPIFLMGNEKMLTKDAKOLEYAAKQFTP 172
|||||
Db 113 viktksiridetitwotikgkpvvyoettdpiflmgnekmltkkdakqleyaakqftp 172
|||||
QY 173 LSLSEFLDRNNTPLWSSRPHNPMYVLPFIMHGKPNRSPNTPSHEARQFTNEFRAPELKF 232
|||||
Db 173 lsisfldrntplwssrphn pmyvlpifmngkpnrspntpsheakqftnefrapelkf 232
|||||
QY 233 QVSVKAAEDLWGTDSLWFGYTQQSHWQIFNGKNSRPRVRHDYQPEIFLTQPVYSDLP 292
|||||
Db 233 qvsvkkaeedlwgtdsdlwfgytqqshwqifngknsrprvrhdypqeilftqpvysdpl 292
|||||
QY 293 WDKVRMIGMGAHVHSGESAKLSRSWNRAYLMAGMEWKNLTVMPRTWGRIFKEGSGSQP 352
|||||
Db 293 wdkvrmi gmga hvhsgesaklsrswnraylmagmewknl tvmprtwgrifkegsgsqp 352
|||||
QY 353 DDNPDI LDYGYGDVRFYOLENKSNTSGTVRYNPRSGKALOLDYVYPLGKISGYFOI 412
|||||
Db 353 ddnpdildygygdvrfyolenksntsgtvrynp rsgkaloldyvyplgk isgyfqi 412
|||||
QY 413 FQYGGSLIDYNHEATSFVGGLMLNDWMGL 442
|||||
Db 413 fqyggslidynheatsfgvglmndwmgl 442
|||||

RESULT 4
AAY85268
ID AAY85268 standard; Protein; 442 AA.
XX
AC AAY85268;
XX
DT 29-JUN-2000 (first entry)
XX
DE BASB034 amino acid sequence #1.
XX
KW Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW hearing loss; antibacterial drug.
XX
OS Moraxella catarrhalis.
XX
PN WO200015802-A1.
XX
PD 23-MAR-2000.
XX
PF 14-SEP-1999; 99WO-EP06781.
XX
PR 14-SEP-1998; 98GB-0020002.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Ruelle J;
XX
DR WPI; 2000-271440/23.
DR N-PSDB; AAA10700.

XX Novel BASB034 polynucleotides and polypeptides from Moraxella
PT catarrhalis used to prepare vaccines against bacterial infections
XX
PS Claim 3; Fig 2; 106pp; English.
XX
CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC strain Mc2931 (ATCC 43617). The invention relates to BASB034 polypeptides
CC from M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The
CC BASB034 polynucleotides and polypeptides may be employed as research
CC reagents and material for the discovery of treatments and diagnostics for
CC diseases, particularly human diseases. They are particularly used to
CC diagnose and treat M. catarrhalis infections. They can be used for
CC diagnosis of disease, staging of disease, or determining response of an
CC infectious organism to drugs. The polynucleotides may be used as a source
CC for hybridization probes, and for screening of genetic mutations,
CC serotype, organism or strain identification, identification of mutations
CC in BASB034 sequences, and as components of arrays which are useful for
CC produce antibodies. The polypeptides can also be used in vaccine
CC formulations, and to identify agonists and antagonists. The polypeptides,
CC antibodies, agonists and antagonists (which are bacteriostatic) are used
CC for the treatment and prevention of diseases such as otitis media in
CC infants and children, pneumonia in elderly, sinusitis, nosocomial
CC infections and invasive diseases, and chronic otitis media with hearing
CC loss. The polypeptides, agonists and antagonists are also used for
CC screening of antibacterial drugs. The BASB034 products of the invention
CC can be used screen for new antibacterial compounds that may target
CC resistant bacteria.
XX
SQ Sequence 442 AA;

Query Match 50.7%; Score 224; DB 21; Length 442;
Best Local Similarity 100.0%; Pred. No. 1.4e-209;
Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 219 QTTNPEFRAPELKFQSVKVAEADLWGTDSLWFGYQQSHWQFNKNSRPFVRHVDYQ 278
DB 219 qttnefrapelkfqsvkvkaeadlwtgdsdlwfgyqqshwqfnknsrpfvrhdyq 278
QY 279 PEIFLTQPVYSLPDWDGVRMIGMCAVHSHNGESAKLSRSNRAYLMAGMEWKNLTVMR 338
DB 279 peifltqpvysldpwdgvrmlgmcavhshngesaklsrsnraylmagmewknlvtmvr 338
QY 339 IWRGIFKEGSGSQPDNDPILDYGYGVDRFLYQLENKSNISGTVRYNPRSGKGLQLDY 398
DB 339 lwrgifkegsgsqpdndpildygygvdrfilyqlenksnlsrgctvrynprrsgkglqldy 398
QY 399 VYPLGKGISGYFQIFQGYGQSLIDYNHEATSGVGLMLNDWML 442
DB 399 vyplgkgisgyfqiifqgygqslidynheatsfgvglmndwml 442

RESULT 5
AAV75156
ID AAV75156 standard; Protein; 370 AA.
XX
AC AAV75156;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria gonorrhoeae ORF 582 protein sequence SEQ ID NO:1786.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy.
XX
OS Neisseria gonorrhoeae.
XX
PN WO9957280-A2.
PD 11-NOV-1999.

XX 30-APR-1999; 99WO-US09346.
PF
XX 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizzo M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
DR WPI: 2000-062150/05.
DR N-PSDB; AAZ53918.
XX
XX Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics -
XX
PS Claim 2; Page 903; 1453pp; English.
CC
CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAZ75941
CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
SQ Sequence 370 AA;
Query Match 2.0%; Score 9; DB 21; Length 370;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 279 PEIFLTQPV 287
DB 209 peifltqpv 217
RESULT 6
AAV75157
ID AAV75157 standard; Protein; 370 AA.
XX
AC AAV75157;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 582 protein sequence SEQ ID NO:1788.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy.
XX
OS Neisseria meningitidis.
XX
PN WO9957280-A2.
XX
PD 11-NOV-1999.

PF 30-APR-1999; 99WO-US09346.
 XX 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.

XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX WPI; 2000-062150/05.
 DR N-PSDB; AA253919.

XX Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics -
 PT
 XX Claim 2; Page 904; 1453pp; English.

XX AA253015 to AA254536, AA254577 to AA254615, and AA254615 to AA254615
 CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA254616 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.

XX Sequence 370 AA;

Query Match 2.0%; Score 9; DB 21; Length 370;
 Best Local Similarity 100.0%; Pred. No. 2.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIFLTQPV 287
 |||||
 Db 209 peifltqpv 217

RESULT 7
 AA253015 standard; Protein; 370 AA.

XX AA253015;
 XX 21-MAR-2000 (first entry)

XX Neisseria meningitidis ORF 582 protein sequence SEQ ID NO:1790.
 DE Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
 KW antibacterial; gene therapy.
 XX Neisseria meningitidis.

XX WO9957280-A2.
 XX 11-NOV-1999.

XX 30-APR-1999; 99WO-US09346.
 PF

XX 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.

XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX WPI; 2000-062150/05.
 DR N-PSDB; AA253920.

XX Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics -
 PT
 XX Claim 2; Page 905; 1453pp; English.

XX AA253015 to AA254536, AA254577 to AA254615, and AA254615 to AA254615
 CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA254616 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.

XX Sequence 370 AA;

Query Match 2.0%; Score 9; DB 21; Length 370;
 Best Local Similarity 100.0%; Pred. No. 2.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIFLTQPV 287
 |||||
 Db 209 peifltqpv 217

RESULT 8
 AA253015 standard; Protein; 374 AA.

XX AA253015;
 XX 18-JUL-2000 (first entry)

XX Neisseria meningitidis serogroup B strain H44/76 BASB033 protein.
 DE Neisseria meningitidis; Neisseria gonorrhoeae; treatment; antibacterial; vaccine;
 KW BASB033; diagnosis, prophylaxis; treatment; antibacterial; vaccine;
 KW Neisseria meningitidis infection.
 XX Neisseria meningitidis.

XX WO200015801-A1.
 XX 23-MAR-2000.

XX 09-SEP-1999; 99WO-EF06718.
 XX 14-SEP-1998; 98GB-0020003.
 PR

XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX Ruelle J;
XX WPI; 2000-271439/23.
XX N-PSDB; AAZ52134.
XX Isolated BASB033 polypeptides and polynucleotides of Neisseria
PT meningitidis, useful for diagnosis, prophylaxis and treatment of N.
PT meningitidis infection -
XX Claim 4; Page 59; 93pp; English.
XX The present sequence is a BASB033 protein from
CC Neisseria meningitidis serogroup B strain H44776. The protein
CC shows homology to the Klebsiella pneumoniae outer membrane
CC phospholipase A. The present sequence is useful for diagnosis,
CC prophylaxis and treatment of N. meningitidis infection. It may also be
CC used for the discovery and development of antibacterial compounds and
CC in vaccine compositions.
XX Sequence 374 AA;
SQ

Query Match 2.0%; Score 9; DB 21; Length 374;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 279 PEIFLTQPV 287
| | | | | | | | | |
DB 213 pelfltqpv 221

RESULT 9
AAY70628
ID AAY70628 standard; Protein; 375 AA.
XX
AC AAY70628;
XX
DT 18-JUL-2000 (first entry)
XX
DE Neisseria meningitidis serogroup B strain ATCC13090 BASB033 protein.
XX
KW BASB033; diagnosis, prophylaxis; treatment; antibacterial; vaccine;
KW Neisseria meningitidis infection.
XX
OS Neisseria meningitidis.
XX
PN WO200015801-A1.
XX
PD 23-MAR-2000.
XX
PF 09-SEP-1999; 99WO-EP06718.
XX
PR 14-SEP-1998; 98GB-0020003.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Ruelle J;
XX
DR WPI; 2000-271439/23.
DR N-PSDB; AAZ52133.
XX
XX Isolated BASB033 polypeptides and polynucleotides of Neisseria
PT meningitidis, useful for diagnosis, prophylaxis and treatment of N.
PT meningitidis infection -
XX Claim 4; Page 58; 93pp; English.
XX The present sequence is a BASB033 protein from
CC Neisseria meningitidis serogroup B strain ATCC13090. The protein
CC shows homology to the Klebsiella pneumoniae outer membrane

CC phospholipase A. The present sequence is useful for diagnosis,
CC prophylaxis and treatment of N. meningitidis infection. It may also be
CC used for the discovery and development of antibacterial compounds and
CC in vaccine compositions.
XX
SQ Sequence 375 AA;
XX

Query Match 2.0%; Score 9; DB 21; Length 375;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 279 PEIFLTQPV 287
| | | | | | | | | |
DB 214 pelfltqpv 222

RESULT 10
AAG11137
ID AAG11137 standard; Protein; 278 AA.
XX
AC AAG11137;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 9741.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127482.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 8; DB 21; Length 278;
Best Local Similarity 100.0%; Pred. No. 16;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 SISFOLDLR 181
 |||||||
Db 271 sisfldlr 278

RESULT 11

AAG11136
ID AAG11136 standard; Protein; 283 AA.

XX
AC AAG11136;

XX
DT 17-OCT-2000 (first entry)

XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 9740.

XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX
OS Arabidopsis thaliana.

XX
PN EP1033405-A2.

XX
PD 06-SEP-2000.

XX
PF 25-FEB-2000; 2000EP-0301439.

XX
PR 25-FEB-1999; 99US-0121825.

PR
PR 05-MAR-1999; 99US-0123180.

PR
PR 09-MAR-1999; 99US-0123548.

PR
PR 23-MAR-1999; 99US-0125788.

PR
PR 25-MAR-1999; 99US-0126264.

PR
PR 29-MAR-1999; 99US-0126785.

PR
PR 01-APR-1999; 99US-0127462.

PR
PR 06-APR-1999; 99US-0128234.

PR
PR 08-APR-1999; 99US-0128714.

PR
PR 16-APR-1999; 99US-0129845.

PR
PR 19-APR-1999; 99US-0130077.

PR
PR 21-APR-1999; 99US-0130449.

PR
PR 23-APR-1999; 99US-0130510.

PR
PR 23-APR-1999; 99US-0130891.

PR
PR 28-APR-1999; 99US-0131449.

PR
PR 30-APR-1999; 99US-0132048.

PR
PR 30-APR-1999; 99US-0132407.

PR
PR 04-MAY-1999; 99US-0132484.

PR
PR 05-MAY-1999; 99US-0132485.

PR
PR 06-MAY-1999; 99US-0132486.

PR
PR 06-MAY-1999; 99US-0132487.

PR
PR 07-MAY-1999; 99US-0132863.

PR
PR 11-MAY-1999; 99US-0134256.

PR
PR 14-MAY-1999; 99US-0134218.

PR
PR 14-MAY-1999; 99US-0134219.

PR
PR 14-MAY-1999; 99US-0134221.

PR
PR 14-MAY-1999; 99US-0134370.

PR
PR 18-MAY-1999; 99US-0134768.

PR
PR 19-MAY-1999; 99US-0134941.

PR
PR 20-MAY-1999; 99US-0135124.

PR
PR 21-MAY-1999; 99US-0135353.

PR
PR 24-MAY-1999; 99US-0135629.

PR
PR 25-MAY-1999; 99US-0136021.

PR
PR 27-MAY-1999; 99US-0136392.

PR
PR 28-MAY-1999; 99US-0136782.

PR
PR 01-JUN-1999; 99US-0137222.

PR
PR 03-JUN-1999; 99US-0137528.

PR
PR 04-JUN-1999; 99US-0137528.

PR
PR 07-JUN-1999; 99US-0137502.

PR
PR 08-JUN-1999; 99US-0137724.

PR
PR 08-JUN-1999; 99US-0138094.

PR
PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.

Query Match 1.88; Score 8; DB 21; Length 283;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 174 SLSFDLDR 181
|||||||

Db 276 slsfdldr 283
RESULT 12
AAG11135
ID AAG11135 standard; Protein; 299 AA.
XX
AC AAG11135;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 9739.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 08-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145216.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147202.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148564.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149155.

PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 8; DB 21; Length 299;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 SLSFDLDR 181
|||||||
Db. 292 slsfdldr 299

RESULT 13
AAR52127
ID AAR52127 standard; Peptide; 16 AA.
XX

```

AC AAR52127;
XX
XX 27-SEP-1996 (first entry)
XX
XX Mouse light chain surface patch S03410.
XX
XX antibody; humanised; murine; heavy chain; light; variable;
KW framework region; complementarity determining region; reshaping;
KW modelling; surface residue; modify.
XX
XX Mus sp.
OS
XX EP592106-A1.
XX
XX 13-APR-1994.
XX
XX 07-SEP-1993; 93EP-0307051.
XX
XX 09-SEP-1992; 92US-0942245.
XX
XX (PEDE/) PEDERSEN J T.
XX (IMMU-) IMMUNOGEN INC.
XX
XX Guild BC, Pedersen JT, Rees AR, Roguska MA, Searle SMJ;
XX WPI; 1994-120230/15.
XX
XX Method of resurfacing of rodent antibodies to produce humanised
XX antibody forms - for producing non-human antibodies with improved
XX therapeutic efficiency by presenting human surface on V-region
XX
XX Example 1; Page 14; 230pp; English.
XX
XX Modification of a rodent antibody or fragment by resurfacing in order
XX to produce a humanised rodent antibody can be determined by calculating
XX homology between murine and human antibody surfaces. In order to test
XX the resurfacing approach of the invention, three humanisation
XX experiments were set up. (1) traditional loop grafting; (2) resurfacing
XX approach using most similar chain; and (3) resurfacing approach using
XX human sequences with most similar surface residues. AAR52069-159 are the
XX surface residue patterns in mouse light chain antibody variable regions.
XX These "patches" were used in the third method, where rodent light and
XX heavy chains were matched and the most similar human sequence found
XX independently only over the surface residues indicated in AAR52030-67.
XX
XX Sequence 16 AA;
SQ
Query Match 1.6%; Score 7; DB 15; Length 16;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 387 PRSGKA 393
Db 4 prsgka 10
|||||

RESULT 14
AAY79483
ID AAY79483 standard; peptide; 50 AA.
XX
XX AAY79483;
XX
XX 01-AUG-2000 (first entry)
XX
XX Rat alpha tubulin variant peptide (aa314-363).
XX
XX Rat; alpha tubulin; mutant; mutein; variant;
XX metabolic X syndrome; hypertension; stroke; diabetes;
KW insulin resistance; obesity; dyslipidemia; anorectic;
KW hypotensive; cerebroprotective; vasotropic; antidiabetic;
KW antilipemic; gene therapy; spontaneously hypertensive; SHR.
XX
XX

```

```

OS Rattus sp.
XX
XX Key Location/Qualifiers
XX Peptide 1..50
XX /note= "corresponds to residues 314-363 of
XX full-length protein"
XX
XX Misc-difference 27
XX /note= "replaces Thr in wild-type"
XX
XX WO200018918-A2.
XX
XX 06-APR-2000.
XX
XX 28-SEP-1999; 99WO-US22494.
XX
XX 28-SEP-1998; 98US-0161939.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shinkets RA;
XX
XX WPI; 2000-303450/26.
XX
XX Novel genes encoding e.g. CD36, SGLT2, and kynurenine aminotransferase,
XX which are predictive and therapeutic for stroke, hypertension, diabetes
XX and obesity
XX
XX Claim 2; Fig 4G; 79pp; English.
XX
XX This sequence represents amino acids 314-363 of alpha-tubulin
XX from a spontaneously hypertensive rat (SHR). This region of
XX alpha-tubulin carries a Ser for Thr amino acid substitution when
XX compared to the corresponding region in control Wistar Kyoto (WKY)
XX rats (see AAY79484). The alpha-tubulin gene is differentially
XX expressed in SHR, SHR-stroke prone (SP) and WKY rats. It may be
XX involved in a predisposition to vascular injury. The SHR rat
XX is an animal model for human metabolic X syndrome. Genes encoding
XX sodium dependent glucose cotransporter, kynurenine aminotransferase,
XX CD36, aldolase A, atrial natriuretic factor, alpha-cardiac myosin
XX and alpha-tubulin were identified as being potentially associated
XX with hypertension, obesity and insulin resistance. These
XX proteins, including isolated proteins comprising the present
XX sequence, are useful for treating, preventing and diagnosing
XX ischemic and metabolic diseases and disorders, such as stroke,
XX hypertension, diabetes and obesity, especially insulin resistivity,
XX dyslipidemia and ischemic stroke (all claimed). Polynucleotides
XX encoding them may also be used in gene therapy and antisense
XX therapy protocols.
XX
XX Sequence 50 AA;
SQ
Query Match 1.6%; Score 7; DB 21; Length 50;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 IKTKRSI 120
Db 22 iktkrsi 28
|||||

RESULT 15
AAG06322
ID AAG06322 standard; Protein; 64 AA.
XX
XX AAG06322;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 3055.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW

```

KW termination sequence.
XX Arabidopsis thaliana.
OS EP1033405-A2.
PN 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 09-MAR-1999; 99US-0123180.
XX 23-MAR-1999; 99US-0123548.
XX 25-MAR-1999; 99US-0125788.
XX 29-MAR-1999; 99US-0126264.
XX 01-APR-1999; 99US-0126785.
XX 06-APR-1999; 99US-0127462.
XX 08-APR-1999; 99US-0128234.
XX 16-APR-1999; 99US-0128714.
XX 19-APR-1999; 99US-0129845.
XX 21-APR-1999; 99US-0130077.
XX 23-APR-1999; 99US-0130449.
XX 28-APR-1999; 99US-0130510.
XX 30-APR-1999; 99US-0130691.
XX 04-MAY-1999; 99US-0131449.
XX 05-MAY-1999; 99US-0132048.
XX 06-MAY-1999; 99US-0132484.
XX 07-MAY-1999; 99US-0132485.
XX 11-MAY-1999; 99US-0132486.
XX 14-MAY-1999; 99US-0132487.
XX 14-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 18-MAY-1999; 99US-0134370.
XX 19-MAY-1999; 99US-0134768.
XX 20-MAY-1999; 99US-0134941.
XX 21-MAY-1999; 99US-0135124.
XX 24-MAY-1999; 99US-0135353.
XX 25-MAY-1999; 99US-0135629.
XX 27-MAY-1999; 99US-0136021.
XX 28-MAY-1999; 99US-0136392.
XX 01-JUN-1999; 99US-0136782.
XX 03-JUN-1999; 99US-0137222.
XX 04-JUN-1999; 99US-0137528.
XX 07-JUN-1999; 99US-0137502.
XX 08-JUN-1999; 99US-0137724.
XX 10-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 14-JUN-1999; 99US-0138847.
XX 16-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 17-JUN-1999; 99US-0139453.
XX 18-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 21-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 24-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140693.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.

PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161320.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.6%; Score 7; DB 21; Length 64;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 162 QLEYAAK 168
Db 44 qleyaak 50

Search completed: May 9, 2002, 12:54:58
Job time: 513 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:55:41 ; Search time 36.87 seconds
(without alignments)
269.771 Million cell updates/sec

Title: US-09-787-083-4

Perfect score: 442

Sequence: 1 MKVSLTLTLILPCFAILA.....YNHEATSFVGLMLNDWMGL 442

Scoring table: OLIGO

Gapex 60.0 , Gapext 60.0

Searched: 212252 seqs, 22503292 residues

Word size : 0

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
1	7	1.6	16	1 US-07-942-245-98
2	7	1.6	552	1 US-08-231-729B-6
3	7	1.6	900	4 US-08-890-865A-4
4	7	1.6	934	1 US-08-215-805A-80
5	7	1.6	1026	2 US-08-542-003-6
6	7	1.6	1026	2 US-08-322-760A-6
7	6	1.4	10	3 US-08-974-775-30
8	6	1.4	11	3 US-08-974-775-29
9	6	1.4	12	3 US-08-974-775-28
10	6	1.4	13	3 US-08-974-775-8
11	6	1.4	13	3 US-08-974-775-27
12	6	1.4	14	3 US-08-974-775-9
13	6	1.4	14	3 US-08-974-775-10
14	6	1.4	14	3 US-08-974-775-32
15	6	1.4	15	3 US-08-974-775-5
16	6	1.4	16	3 US-08-974-775-4
17	6	1.4	20	1 US-08-430-273-1
18	6	1.4	21	1 US-07-918-953-13
19	6	1.4	21	1 US-07-918-953-15
20	6	1.4	21	1 US-08-212-696-1
21	6	1.4	21	1 US-08-158-245-1
22	6	1.4	21	1 US-08-081-661-13
23	6	1.4	21	1 US-08-081-661-15
24	6	1.4	21	1 US-08-233-617-1
25	6	1.4	21	1 US-08-160-376A-1
26	6	1.4	21	1 US-08-304-070-1
27	6	1.4	21	1 US-08-285-661-1

28	6	1.4	21	1	US-08-301-838-1	Sequence 1, Appli
29	6	1.4	21	1	US-08-389-487-4	Sequence 4, Appli
30	6	1.4	21	1	US-08-389-487-12	Sequence 12, Appli
31	6	1.4	21	1	US-08-507-124-2	Sequence 2, Appli
32	6	1.4	21	1	US-08-507-124-3	Sequence 3, Appli
33	6	1.4	21	1	US-08-342-931-1	Sequence 1, Appli
34	6	1.4	21	1	US-08-400-256-1	Sequence 1, Appli
35	6	1.4	21	2	US-08-508-664-9	Sequence 9, Appli
36	6	1.4	21	2	US-08-353-476-85	Sequence 85, Appli
37	6	1.4	21	2	US-08-353-476-87	Sequence 87, Appli
38	6	1.4	21	2	US-08-484-219-1	Sequence 1, Appli
39	6	1.4	21	2	US-08-979-587-1	Sequence 1, Appli
40	6	1.4	21	2	US-08-992-676-1	Sequence 1, Appli
41	6	1.4	21	2	US-08-992-676-6	Sequence 6, Appli
42	6	1.4	21	2	US-09-134-836-1	Sequence 1, Appli
43	6	1.4	21	2	US-09-134-836-7	Sequence 7, Appli
44	6	1.4	21	3	US-08-967-867-2	Sequence 2, Appli
45	6	1.4	21	4	US-08-932-082-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-07-942-245-98
; Sequence 98, Application US/07942245
; Patent No. 5639641
; GENERAL INFORMATION:
; APPLICANT: PEDERSEN, Jan T.
; APPLICANT: SEARLE, Stephen M.J.
; APPLICANT: REES, Anthony R.
; APPLICANT: ROGUSKA, Michael A.
; APPLICANT: GUILD, Braydon C.
; TITLE OF INVENTION: SURFACE RESIDUE VENEERING OF RODENT
; TITLE OF INVENTION: ANTIBODIES
; NUMBER OF SEQUENCES: 522
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughree, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: HP 9000/700 Workstation
; OPERATING SYSTEM: UNIX
; SOFTWARE: In house
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07942,245
; FILING DATE: 09-SEP-1992
; CLASSIFICATION: 530
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-942-245-98

Query Match 1.6%; Score 7; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 387 PRSGKA 393
Db 4 PRSGKA 10

```
RESULT 2
US-08-231-729B-6
; Sequence 6, Application US/08231729B
; Patent No. 5618722
; GENERAL INFORMATION:
; APPLICANT: ZENNO, Shuhei
; APPLICANT: SHIRAIISHI, Shinji
; APPLICANT: INOUE, Satoshi
; APPLICANT: SAIGO, Kaoru
; TITLE OF INVENTION: FIREFLY LUCIFERASE GENE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LEYDIG, VOIT & MAYER
; STREET: 700 Thirteenth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,729B
; FILING DATE: 20-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 119050/1993
; FILING DATE: 21-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rose, Herbert C.
; REGISTRATION NUMBER: 29846
; REFERENCE/DOCKET NUMBER: 60130/No. 5618722aka
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-737-6770
; TELEFAX: 202-737-6776
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 552 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-231-729B-6

Query Match 1.6%; Score 7; DB 1; Length 552;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 LSFDLDR 181
Db 184 LSFDLDR 190

RESULT 3
US-08-890-865A-4
; Sequence 4, Application US/08890865A
; Patent No. 6307019
; GENERAL INFORMATION:
; APPLICANT: Constantini, Franklin
; APPLICANT: Zeng, Li
; TITLE OF INVENTION: AXIN GENE AND USES THEREOF
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10036
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/890,865A
; FILING DATE: 10-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/54249
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)278-0400
; TELEFAX: (212)391-0526
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 900 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-890-865A-4

Query Match 1.6%; Score 7; DB 4; Length 900;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 403 GKGISGY 409
Db 268 GKGISGY 274

RESULT 4
US-08-215-805A-80
; Sequence 80, Application US/08215805A
; Patent No. 5559008
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; TITLE OF INVENTION: LEUKOTOXIN GENE FROM PASTURELLA
; NUMBER OF SEQUENCES: 84
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/215,805A
; FILING DATE: 22-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Timian, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 19603/61 (D-1329A)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1636
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 80:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 934 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```

```
; ORIGINAL SOURCE:
; ORGANISM: Pasteurella suis
; STRAIN: 5943
; IMMEDIATE SOURCE:
; LIBRARY: P. suis DNA in Bacteriophage lambda-dash
; CLONE: (Lambda)yfc33-37
; US-08-215-805A-80

Query Match          1.6%; Score 7; DB 1; Length 934;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 238 VKAEDL 244
Db 62 VKAEDL 68

RESULT 5
US-08-542-003-6
; Sequence 6, Application US/08542003
; Patent No. 5864013
; GENERAL INFORMATION:
; APPLICANT: Goldberg, Edward B.
; TITLE OF INVENTION: MATERIALS FOR THE PRODUCTION OF
; TITLE OF INVENTION: NANOMETER STRUCTURES AND USE THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie and Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/542,003
; FILING DATE: 13-OCT-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 8471-0005-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: 212-869-8864
; TELEX: 66441 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1026 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Bacteriophage T4
; IMMEDIATE SOURCE:
; CLONE: p37 amino acid
; US-08-542-003-6

Query Match          1.6%; Score 7; DB 2; Length 1026;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 QTIGKP 134
Db 861 QTIGKP 867
```

```
RESULT 6
US-08-322-760A-6
; Sequence 6, Application US/08322760A
; Patent No. 5877279
; GENERAL INFORMATION:
; APPLICANT: Goldberg, Edward B.
; TITLE OF INVENTION: MATERIALS FOR THE PRODUCTION OF
; TITLE OF INVENTION: NANOMETER STRUCTURES AND USE THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie and Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/322,760A
; FILING DATE: 13-OCT-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 8471-0003-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: 212-869-8864
; TELEX: 66441 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1026 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Bacteriophage T4
; IMMEDIATE SOURCE:
; CLONE: p37 amino acid
; US-08-322-760A-6

Query Match          1.6%; Score 7; DB 2; Length 1026;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 QTIGKP 134
Db 861 QTIGKP 867

RESULT 7
US-08-974-775-30
; Sequence 30, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
```

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-974-775-30

Query Match 1.4%; Score 6; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 424 NHEATS 429
Db 5 NHEATS 10

RESULT 8
US-08-974-775-29
; Sequence 29, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; STREET: Plaza Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
```

```
;
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-974-775-29

Query Match 1.4%; Score 6; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 424 NHEATS 429
Db 5 NHEATS 10

RESULT 9
US-08-974-775-28
; Sequence 28, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; STREET: Plaza Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-974-775-28

Query Match 1.4%; Score 6; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 424 NHEATS 429
Db 5 NHEATS 10

RESULT 10
US-08-974-775-8
; Sequence 8, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
```

APPLICANT: Toback, F. Gary
APPLICANT: Walsh-Reitz, Margaret
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
STREET: Plaza Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775
FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-8

Query Match 1.4%; Score 6; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
|||||
DB 8 NHEATS 13

RESULT 11
US-08-974-775-27
Sequence 27, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:
APPLICANT: Toback, F. Gary
APPLICANT: Walsh-Reitz, Margaret
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
STREET: Plaza Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775

FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-27

Query Match 1.4%; Score 6; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
|||||
DB 5 NHEATS 10

RESULT 12
US-08-974-775-9
Sequence 9, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:
APPLICANT: Toback, F. Gary
APPLICANT: Walsh-Reitz, Margaret
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
STREET: Plaza Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775
FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-9

Query Match 1.4%; Score 6; DB 3; Length 14;

Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
|||||
DB 8 NHEATS 13

RESULT 13

US-08-974-775-10
; Sequence 10, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-974-775-10

Query Match 1.4%; Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
|||||
DB 8 NHEATS 13

RESULT 14

US-08-974-775-32
; Sequence 32, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE

; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 14
; OTHER INFORMATION: /product= "Ser or Tyr"
US-08-974-775-32

Query Match 1.4%; Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
|||||
DB 8 NHEATS 13

RESULT 15

US-08-974-775-5
; Sequence 5, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:


```

; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-974-775-5

```

```

Query Match      1.4%; Score 6; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
    |||||
Db 8 NHEATS 13

```

Search completed: May 9, 2002, 12:55:42
Job time: 552 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:14:18 ; Search time 44.45 Seconds
(without alignments)
757.461 Million cell updates/sec

Title: US-09-787-083-2
Perfect score: 2360
Sequence: 1 MKVSLSTLTSLSCFAILA.....YNHEATSRFGVGLMLNDWMGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	814	34.5	382	E81195	phospholipase A1,
2	814	34.5	409	H81831	probable phospholip
3	373	15.8	286	B36971	outer membrane pho
4	367	15.6	329	D81279	phospholipase A1 (
5	358	15.2	289	A36971	outer membrane pho
6	357	15.1	289	P5E2A1	phospholipase A1 (
7	357	15.1	289	E86059	outer membrane pho
8	321.5	13.6	289	C36971	outer membrane pho
9	246.5	10.4	355	C64582	phospholipase A1 p
10	240.5	10.2	355	H71930	probable phospholi
11	112.5	4.8	602	TVRTRR	protein kinase (EC
12	110.5	4.7	800	A29003	cellulase (EC 3.2.
13	110.5	4.7	822	JT0611	cellulase (EC 3.2.
14	109	4.6	824	JC7532	cellulase (EC 3.2.
15	107.5	4.6	901	T20122	hypothetical prote
16	104	4.4	783	JC5467	cellulase (EC 3.2.
17	101.5	4.3	660	E83656	methionyl-tRNA syn
18	101.5	4.3	838	A96557	probable receptor
19	100.5	4.3	1658	T42642	phosphoinositide 3
20	99	4.2	5005	F82884	hypothetical prote
21	98.5	4.2	797	T46737	X-Pro dipeptidyl-p
22	98.5	4.2	810	S49744	AMP deaminase (EC
23	97	4.1	4307	T20721	hypothetical prote
24	96	4.1	719	A42893	penicillin-binding
25	96	4.1	888	S50801	AMP deaminase homo
26	95	4.0	751	H96839	hypothetical prote
27	94.5	4.0	397	A35136	cellulase (EC 3.2.
28	94.5	4.0	564	T40777	ferric reductase t
29	94.5	4.0	601	T26062	hypothetical prote

30 94.5 4.0 655 2 T26061 hypothetical prote
31 94.5 4.0 1310 2 T40135 probable involveme
32 94 4.0 719 2 S28031 penicillin-binding
33 94 4.0 719 2 S28033 penicillin-binding
34 94 4.0 765 2 T35719 chitinase - Strept
35 94 4.0 1078 2 T18352 protein p120 - Myc
36 93.5 4.0 1478 2 S20117 protein kinase BCK
37 93 3.9 719 2 S28034 penicillin-binding
38 93 3.9 719 2 S28032 penicillin-binding
39 92.5 3.9 857 1 A41369 S-receptor kinase
40 92 3.9 324 2 B69521 hypothetical prote
41 92 3.9 470 2 T43675 cog-2 protein - Ca
42 92 3.9 1788 2 T29043 hypothetical prote
43 91.5 3.9 467 1 A49377 involucrin - mouse
44 91.5 3.9 486 2 S30959 gene 14 protein -
45 91.5 3.9 711 2 E71673 hypothetical prote

ALIGNMENTS

RESULT 1

E81195

phospholipase A1, probable NMB0464 [imported] - Neisseria meningitidis (strain MC58 s
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: E81195
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizzia, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755
A;Accession: E81195
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-382 <TET>
A;Cross-references: GB:AE002403; GB:AE002098; NID:g7225688; PIDN:AAF40901.1; PID:g722
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0464

Query Match	34.5%	Score 814	DB 2	Length 382
Best Local Similarity	44.4%	Pred. No. 3.6e-58		
Matches 162	Conservative	68	Mismatches 111	Indels 24
Gaps				8
Qy	87	INCSALNODIMRLACYDTLVHGETPAVI-----RTKRSIRLDETWTQTI-RGKQVVIYQE	140	
Db	33	LQCAALTDNVTRLACYDRIFAAQLPSSAGQESQESKAVLNLTETVRSLLDKGEAVIVVEK	92	
Qy	141	TTDPFILMGNEKGMILTKDAKQLEYAAKQFTPLSLFSDLDNRN-TPLWSSRPHNPMYVLP	199	
Db	93	GGDAL-----PADSAGETADITYPLSLMTYDLKNDLRLGLLVGRHNPMLMP	139	
Qy	200	IFMHGKNRSPNTPSH-EAKQFTNEFRAPELKFQVSVKVAEDLWCTDSDLNFGYTTQ	258	
Db	140	LWYNNSNPAPGSPTRGTTVOEKFGQQRRAETKLVQSPKSKIAEDLFTRADLFGYITQR	199	
Qy	259	SHWQIFN-GKNSRPFRRVHDYQPEIFLTQPVYSDLPWDCVKRMIGMGAHVHSHNGESAKLSR	317	
Db	200	SDWQIYNOGRKSAPFRNTDYKPEIFLTQPVKADLPFGRLRLMGLGAFVHQSGQSRPSR	259	
Qy	318	SWNRAYLMAGMEWKNLTVMPIRWIRIFKSGSQPDNDPDLIDYVYGVDVRFPLENKS	377	
Db	260	SWNRAYLMAGMEWKNLTVMPIRWIRIFKSGSQPDNDPDLIDYVYGVDVRFPLENKS	317	
Qy	378	NISVTYRNPRSGKALQLDVYVPLGKIGISGYFQFCYQGLSDIDYNHEATSFVGLMLN	437	
Db	318	NYSVLYRNPKTGYGATEAAAYTFPIKGLKGVGRFGHGBESLIDYNHKNQNGIGIGLMFN	377	
Qy	438	DWMGL 442		

Db 378 DLDGI 382

RESULT
H81831

probable phropholipase NMA2021 [Imported] - Neisseria meningitidis (strain Z2491 serogroup C)
 C:Species: Neisseria meningitidis
 C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
 C:Accession: H81831
 R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
 ; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
 Nature 404, 502-506, 2000
 A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
 A:Reference number: A81775; MUID:20222556
 A:Accession: H81831
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-409 <PAR>
 A:Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85240.1; PID:g738065
 A:Experimental source: serogroup A, strain Z2491
 C:Genetics:
 A:Gene: NMA2021

Query Match	34.5%	Score 814;	DB 2;	Length 409;
Best Local Similarity	44.4%;	Pred. No. 4e-58;		
Matches 162;	Conservative	68;	Mismatches 111;	Indels 24; Gaps 8;

Qy	87	INCSALNQDITRLRACVDTLYVHGETTTPAVI-----KTVRSIRLDETIWQTT-KGKPQIVTQE	140
Dd	60	LQCAALTNDNTYTRILLACY :	119
Qy	141	TTDPIFLMGNEKMLTKKDAKLEYAAKOFTPLSLSFSDLDRNN-TPLWSSSRPHNPMVYL	199
Dd	120	GGDAL-----PADSAGETADITYTPUSLMYDLDDKLNRGLLGVRGHENPMYLP	166
Qy	200	IFMHGKPNRSPNTPSH-EAKQFPTNPFRAPELKFOVSVKVAEEDLWG'TDSOLWFQYTOQ	258
Dd	167	LWYNNSPYAPGSPTRGTTTVEQXFGOOKRAETKLQVSEFKSKIAEDLFKTADLWFGYTQR	226
Qy	259	SHWOIFN-GKNRSPRRVHDYQPIFITQPVYSDDLWDGWKVMIGMGAHVHHNGESAKULSR	317
Dd	227	SDWOIYNQGRKSAPPFRNTDYKPEIFITQPVKADLPFGGBLRMLGALGFVHQSGQSQRPESR	286
Qy	318	SNWRAYLMAGMEWNKLTVMPRINGRIKFEGSGSDPDNDPILDYCYGYGDVRFYQLLENKS	377
Dd	287	SNWRIYAMAGMEWKLTVPJRWVRAFQ-SGBK-NDNPDIADIMGYGVDKLYQJRLNDRQ	344
Qy	378	NISGTVRYNPSRKGLQLLDVYPVLKGIGISGYQIFQYCGSLIDYNNHEATSFYGVGLMLN	437
Dd	345	NVYSVLRYNPKTYGATEAAYTPIKGLKLVVYRGPHGYGESLIDYNNHKONGIGTGICLMFN	404
Qy	438	DWMGL 442	
Dd	405	DLDCI 409	

RESULT
B36971
3

outer membrane phospholipase A (EC 3.1.1.-) precursor - Klebsiella pneumoniae
C/Species: Klebsiella pneumoniae
C/Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 18-Jun-1999
C/Accession: B36971; S40129
R/BROK, R.G.P.M.; Brinkman, E.; van Boxtel, R.; Bekkers, A.C.A.P.A.; Verheij, H.M.; Tomm
J. Bacteriol. 176, 861-870, 1994
A/Title: Molecular characterization of enterobacterial *pldA* genes encoding outer membran
A/Reference number: A36971; MUID:94131966
A/Accession: B36971
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-286 <BRO>
A/Cross-references: EMBL:X76901; NID:g436880; PIDN:CAA54223.1; PID:g436881

A;Note: authors translated the codon AAG for residue 112 as Arg
C;Genetics:
A;Gene: pldA
C;Superfamily: bacterial phospholipase A1
C;Keywords: carboxylic ester hydrolase

Query Match	15.8%	Score	373;	DB 2;	Length	286;			
Best Local Similarity	42.1%;	Pred. NO.	1.2e-22;						
Matches	90;	Conservative	27;	Mismatches	87;	Indels	10;	Gaps	5;

[illegible]

RESULT 4

D81279
phospholipase A1 (EC 3.1.1.32) Cj1351 [Imported] - Campylobacter jejuni (strain NCTC
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 18-Aug-2000
C:Accession: D81279
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth,
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrell,
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals
A:Reference number: AB1250; PMID:20150912
A:Accession: D81279
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-329 <PAR>
A:Cross-references: GB:AL111168; NTID:g6968723; PIDN:CAB73778.1; PID:g6968723
A:Experimental source: serotype OZ, strain NCTC 11168
C:Genetics:
A:Gene: pldA; Cj1351
C:Keywords: carboxylic ester hydrolase

Query Match	15.6%	Score	367;	DB 2;	Length	329;
Best Local Similarity	33.5%	Pred. No.	4.4e-22;			
Matches 106;	Conservative	47;	Mismatches	125;	Indels	38;
Gaps	13;					

[illegible]

QY 422 DYNHEATSFVGLMLN 437
| | | | |
Db 313 DYNKHLQLSTGFLIS 328
| | | | |

RESULT 5
A36971
outer membrane phospholipase A (EC 3.1.1.-) precursor - Salmonella typhimurium
C:Species: Salmonella typhimurium
C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 18-Jun-1999
C:Accession: A36971; S40131
R:Brok, R.G.P.M.; Brinkman, E.; van Bortel, R.; Bekkers, A.C.A.P.A.; Verheij, H.M.; Tomm
J. Bacteriol. 176, 861-870, 1994
A:Title: Molecular characterization of enterobacterial pida genes encoding outer membra
A:Reference number: A36971; MUID:94131966
A:Accession: A36971
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-289 <BRO>
A:Cross-references: GB:X76900; NID:g437024; PIDN:CAA54222.1; PID:g437025
C:Genetics:
A:Gene: pida
A:Superfamily: bacterial phospholipase A1
C:Keywords: carboxylic ester hydrolase

Query Match 15.2%; Score 358; DB 2; Length 289;
Best Local Similarity 39.3%; Pred. No. 2e-21;
Matches 84; Conservative 30; Mismatches 90; Indels 10; Gaps 4;

QY 226 RAPELKFQVSVKAAEDLWGTSDFMGFYTQOSHWOIFNGKNSRPRVHDYQPEIFLTQ 285
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 83 RKDEVKFLQSLAFLPLWRGILGNPSVLGASVYTKSWQLSNKSSESPRETNYEPQLFLG 142
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

QY 286 PV-YSDLPLWDGKVRMIGMGAVHHSNGESAKLSRSNRAYLMAGMEWKNLTVMPIWGRI 344
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

Db 143 ATDYRFAGW--TLRHVEMGYNHDNSGRSDPTSRTNRLYRLMAENGWNLVEVKPWVI- 199
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

QY 345 KEGSGSQDDNPDLTDYGYGDRVFLYOLENKNISGTVRNPNSGKGAQLDYVPLGK 404
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

Db 200 -----GSTDNPDITKMGYQLKIGYHL-GEAVLSAKGQYNWNTGYGGAELGSLYPVT 253
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

QY 405 GISGYFOIFGYGQSLDYNHEATSFVGLMLND 438
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

Db 254 HVRLYTQVSYGSGESLDYFNQTRGVGVMLND 287
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

RESULT 6
PSECA1
phospholipase A1 (EC 3.1.1.32) precursor - Escherichia coli
N:Alternate names: outer-membrane phospholipase A; phosphatidylcholine 1-acylhydrolase;
C:Species: Escherichia coli
C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 11-Jun-1999
C:Accession: A22133; S30711; A00771; S66447; F65186
R:Homma, H.; Kobayashi, T.; Chiba, N.; Karasawa, K.; Mizushima, H.; Kudo, I.; Inoue, K.;
J. Biochem. 96, 1655-1664, 1994
A:Title: The DNA sequence encoding pida gene, the structural gene for detergent-resistan
A:Reference number: A22133; MUID:85157492
A:Accession: A22133
A:Molecule type: DNA
A:Residues: 1-289 <HOM>
A:Cross-references: GB:X02143; GB:X00780; NID:g42423; PIDN:CAA26081.1; PID:g757840
A:Note: this enzyme is tightly bound to the outer membrane of the cell
R:Daniels, D.L.; Plunkett III, G.; Burland, V.; Blattner, F.R.
Science 257, 771-778, 1992
A:Title: Analysis of the Escherichia coli genome: DNA sequence of the region from 84.5 b
A:Reference number: S30660; MUID:92358234
A:Accession: S30711
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-13, 'FA', 16-289 <DAN>
A:Cross-references: EMBL:M87049; NID:g836656; PIDN:AAA67617.1; PID:g148220

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1992
R:de Geus, P.; Verheij, H.M.; Riegman, N.H.; Hoekstra, W.P.M.; de Haas, G.H.
EMBO J. 3, 1799-1802, 1984
A:Title: The pro- and mature forms of the E. coli K-12 outer membrane phospholipase A
A:Reference number: A00771; MUID:85003590
A:Accession: A00771
A:Molecule type: DNA
A:Residues: 'MTRQ', 34-289 <DEG>
R:Dekker, N.; Merck, K.; Tommassen, J.; Verheij, H.M.
Eur. J. Biochem. 232, 214-219, 1995
A:Title: In vitro folding of Escherichia coli outer-membrane phospholipase A.
A:Reference number: S66447; MUID:96048049
A:Accession: S66447
A:Status: preliminary
A:Molecule type: protein
A:Residues: 21-24 <DEK>
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: F65186
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-289 <BLAT>
A:Cross-references: GB:AE000458; GB:U00096; NID:g2367299; PIDN:AAC76824.1; PID:g23673
A:Experimental source: strain K-12, substrain MG1655
C:Comment: This enzyme is tightly bound to the outer membrane of the cell.
C:Genetics:
A:Gene: pida
A:Map position: 85 min
C:Superfamily: bacterial phospholipase A1
C:Keywords: carboxylic ester hydrolase; membrane bound
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-289/Product: phospholipase A1 #status predicted <MPT>

Query Match 15.1%; Score 357; DB 1; Length 289;
Best Local Similarity 39.3%; Pred. No. 2.4e-21;
Matches 84; Conservative 29; Mismatches 91; Indels 10; Gaps 4;

QY 226 RAPELKFQVSVKAAEDLWGTSDFMGFYTQOSHWOIFNGKNSRPRVHDYQPEIFLTQ 285
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

Db 83 RKDEVKFLQSLAFLPLWRGILGNPSVLGASVYTKSWQLSNKSSESPRETNYEPQLFLG 142
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

QY 286 PV-YSDLPLWDGKVRMIGMGAVHHSNGESAKLSRSNRAYLMAGMEWKNLTVMPIWGRI 344
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

Db 143 ATDYRFAGW--TLRDVEMGYNHDNSGRSDPTSRSNRLYRLMAENGWNLVEVKPWVYV- 199
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

QY 345 KEGSGSQDDNPDLTDYGYGDRVFLYOLENKNISGTVRNPNSGKGAQLDYVPLGK 404
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

Db 200 -----GWTDNDPDITKMGYQLKIGYHL-GDAVLSAKGQYNWNTGYGGAELGSLYPITK 253
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

QY 405 GISGYFOIFGYGQSLDYNHEATSFVGLMLND 438
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

Db 254 HVRLYTQVSYGSGESLDYFNQTRGVGVMLND 287
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

RESULT 7
E86069
outer membrane phospholipase A [imported] - Escherichia coli (strain O157:H7)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001
C:Accession: E86069
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: E86069
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-289 <STO>

A:Cross-references: GB:AE005174; NID:gl2518695; PIDN:AAG59017.1; GSPDB:GN00145; UWGP:Z53
A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: pldA

C:Superfamily: bacterial phospholipase A1

Query Match 15.1%; Score 357; DB 2; Length 289;

Best Local Similarity 39.3%; Pred. No. 2.4e-21;

Matches 84; Conservative 29; Mismatches 91; Indels 10; Gaps 4;

QY 226 RAPELKFQVSVKVAEADLWGTDSDLWFGYTOQSHWQIFNGKNSRPFVRVHDYQPEIFLTQ 285

DB 83 RKDEKVFQSLAPFLWRGILGPNVLSGASYTQKSNWQLSNSESPFRETNYEPQLFLGF 142

QY 285 PV-YSDLPWDGKVRMIGCAVHSHNGESAKLSRSNRAYLMAGMEKKNLTVMPTWGRIF 344

DB 143 ATDYRFAGH--TLRDVEMGYNHDNSGRSDPTSRSNRLYTRLMAENGWNLVEVKPMYVY- 199

QY 345 KEGSGSQPDNDPDIIDYGYGVDRFLYOLENKSNTSGTVRYNPRSGKGAQLQLDYVPLGK 404

DB 200 -----GNTDDNDPIYKMGYQKLTGYHL-GDAVLSAKQYNNWTGYGGAELGLSYPTK 253

QY 405 GTSYGFQIFQYQSGSLIDYNHEATSGFVGLMLND 438

DB 254 HVRLTYQVYSGYGESLIDYNFQNTQVGVGLMLND 287

RESULT 8

C36971

outer membrane phospholipase A (EC 3.1.1.1-) precursor - Proteus vulgaris

C:Species: Proteus vulgaris

C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 18-Jun-1999

C:Accession: C36971; 540130

R:Brok, R.G.P.M.; Brinkman, E.; van Bortel, R.; Bekkers, A.C.A.P.A.; Verheij, H.M.; Tomm

J. Bacteriol. 176, 861-870, 1994

A:Title: Molecular characterization of enterobacterial pldA genes encoding outer membra

A:Reference number: A36971; MUID:94131966

A:Accession: C36971

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-289 <BRO>

A:Cross-references: EMBL:X76902; NID:9436889; PIDN:CAA54224.1; PID:9436890

A:Note: authors translated the codon GAG for residue 74 as Gly and CGA for residue 115 a

C:Genetics:

A:Gene: pldA

C:Superfamily: bacterial phospholipase A1

C:Keywords: carboxylic ester hydrolase

Query Match 13.6%; Score 321.5; DB 2; Length 289;

Best Local Similarity 33.3%; Pred. No. 1.8e-18;

Matches 87; Conservative 42; Mismatches 115; Indels 17; Gaps 7;

QY 179 LDRNNTPLSSRPHNPMYVLPFIEMHGKPNRSPNTPSHEAKQFTPNFRAPELKFQVSVKV 238

DB 43 LQEHNDNP-FTLYPYESNYLLTYV-----TSLNKKAIYESYNNSDNA-NKDEVKFKQLSLAF 95

QY 239 KAAEDLWGTDSDLWFGYTOQSHWQIFNGKNSRPFVRVHDYQPEIFLTQPV-YSDLPWDGKV 297

DB 96 PLWRGILGPNLSLGASYTQSRWQLSNTGESAPFRETNYEPQLFLGFATDYSVGDW--TL 153

QY 298 RMIGMCAVHSHNGESAKLSRSNRAYLMAGMEKKNLTVMPTWGRIFKESGSGSQPDNDP 357

DB 154 RAEFGYNHQSNGRSDPTSRSNRLYSRMAONGWNLVEVKPMYVY-----GDTSDNKN 207

QY 358 ILDYYGYGVDRFLYOLENKSNTSGTVRYNPRSGKGAQLQLDYVPLGKISGYFQIFQYQ 417

DB 208 ITKMYGYQKLTGYQL-GEAVLSAKQYNNWTGYGGAELGVSYPITKHFRTYQVYSGYG 266

QY 418 QSLIDYNHEATSGFVGLMLND 438

DB 267 ESLIDYDFNQTRVGMVGLMLND 287

RESULT 9

C64582

phospholipase A1 precursor - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori

C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999

C:Accession: C64582

R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKe

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey,

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,

A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467

A:Accession: C64582

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-355 <TOM>

A:Cross-references: GB:AE000564; GB:AE000511; NID:g2313602; PIDN:AAD07564.1; PID:g231

Query Match 10.4%; Score 246.5; DB 2; Length 355;

Best Local Similarity 25.1%; Pred. No. 2.8e-12;

Matches 82; Conservative 41; Mismatches 117; Indels 87; Gaps 11;

QY 157 KDKAKQLEYAAKQFTPLSLSEFDLDRNNTPLWSSRPHNPMYVLPFIEMHGKPNRSPNTPSHE 216

DB 69 KKYLNMDYLGTYFLPFFYHSF-----TPIFQWYHPNINP----- 102

QY 217 AKQFTPNFRAPELKFQVSVKVAEADLWGTDSDLWFGYTOQSHWQIFNGKNSRPFVRVHD 276

DB 103 ---YQNEF-----KQISFRVPVFRHILWTGTLVAYTQTDWFOIYNDPQSPHRMNM 154

QY 277 YQPEIFLTQPVYSDLPWDGKV---RMIGMCAVHSHNG-ESAKLSRSNRAYLMAGMEKWN 332

DB 155 FMPLEIYVYVPI-NFKPFGKIGNFSEIWIWQHOISNGVGGVGAOCYQPFNK-----EGNPENQ 209

QY 333 LTVMPRI-----WGRIFKEGSGSQP-----DDNDP 357

DB 210 FPGQPVIVDYNGQKDVRRMGGRSVSAGORPFRVLWEKGGGLKIMVAYWYVYDSDNPN 269

QY 358 ILDYYGYGVDRFLY-----QLENKSNTSGTVRYNPRSGKGAQLQLDYVPLGKISGYF 410

DB 270 LIDYMGYGNKIDYRRGRHHEFLQLYDIPTQWRYD--RWHGAFRLGTYTYRINPFVGIYA 327

QY 411 QIFQGYGQSLIDYNHEATSGFVGLMLN 437

DB 328 QWFMNGYGDGLYEYDFNFSNRIGVGIRLN 354

RESULT 10

H71930

probable phospholipase a1 - Helicobacter pylori (strain J99)

C:Species: Helicobacter pylori

A:Variety: strain J99

C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999

C:Accession: H71930

R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.

Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F

Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric p

A:Reference number: A71800; MUID:99120557

A:Accession: H71930

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-355 <ARN>

A:Cross-references: GB:AE001479; GB:AE001439; NID:g4154979; PIDN:AAD06029.1; PID:g415

A:Experimental source: strain J99

C:Genetics:

A:Gene: pldA

[illegible]

Search completed: November 30, 2001, 14:18:32
Job time: 254 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:17:53 ; Search time 28.02 Seconds
(without alignments)
578.367 Million cell updates/sec

Title: US-09-787-083-2
Perfect score: 2360
Sequence: 1 MKVSLSTLTLTILSCFAILA.....YNHEATSFVGLMLNDWMGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 segs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	373	15.8	286	1 PAL_KLEPN	P37446 klebsiella
2	358	15.2	289	1 PAL_SALTY	P37442 salmonella
3	357	15.1	289	1 PAL_ECOLI	P00631 escherichia
4	321.5	13.6	289	1 PAL_PROVU	P37447 proteus vul
5	110.5	4.7	800	1 GUN_BACS1	P06564 bacillus sp
6	98.5	4.2	810	1 AMDM_YEAST	P15274 saccharomyc
7	96	4.1	888	1 YJHO_YEAST	P40361 saccharomyc
8	94.5	4.0	397	1 GUN_PAEPO	P23548 paenibacill
9	93.5	4.0	1478	1 BCKI_YEAST	Q01389 saccharomyc
10	93	3.9	716	1 BAC2_MOUSE	P97303 mus musculu
11	92.5	3.9	849	1 SRK6_BRAOL	Q09092 brassica ol
12	92	3.9	324	1 YL70_ARCFU	Q28112 archaeoglob
13	92	3.9	1158	1 RI14_HUMAN	P48552 homo sapien
14	92	3.9	1788	1 IP72_CAEEL	Q09221 caenorhabdi
15	91.5	3.9	467	1 INVO_MOUSE	P48997 mus musculu
16	91.5	3.9	486	1 VG14_BPML5	Q05220 mycobacteri
17	91.5	3.9	1024	1 Y075_MVCGE	P47321 mycoplasma
18	90.5	3.8	379	1 XJHT_HAEIN	P44544 haemophilus
19	90	3.8	496	1 CATA_DICDI	O77229 dictyosteli
20	90	3.8	503	1 BPBA_STRPN	Q04707 streptococc
21	89.5	3.8	503	1 CD44_RAT	P26051 rattus norv
22	89.5	3.8	825	1 GUN3_BACS4	P19570 bacillus sp
23	89	3.8	521	1 NPKE_BACAM	P06832 bacillus am
24	88.5	3.8	1379	1 MET_MOUSE	P16056 mus musculu
25	88	3.7	1131	1 PMAL_DUNBI	P54211 dunaliella
26	87.5	3.7	353	1 DCUP_BACSU	P32395 bacillus su
27	87.5	3.7	478	1 DHGB_ACTICA	P13650 acinetobact
28	87.5	3.7	669	1 AMY_ALTHA	P29957 alteromonas
29	87.5	3.7	794	1 SEIL_HUMAN	Q9ubv2 homo sapien
30	87.5	3.7	969	1 SACE_STRSL	Q55242 streptococc
31	87	3.7	483	1 MORE_CHLTR	O84271 chlamydia t
32	87	3.7	537	1 P4H2_MOUSE	O60716 mus musculu
33	87	3.7	560	1 DTXH_CORBE	P00589 corynephege

34	87	3.7	842	1 LPFC_SALTY	P43662 salmonella
35	87	3.7	1115	1 DP3A_BACSU	O34623 bacillus su
36	87	3.7	1474	1 A2MG_HUMAN	P01023 homo sapien
37	86.5	3.7	1374	1 YC9A_SCHPO	Q09884 schizosacch
38	86	3.6	491	1 TY3H_PHASP	P11982 phasianidae
39	86	3.6	666	1 PD14_RAT	O88807 rattus norv
40	86	3.6	828	1 BGAL_BRAOL	P49676 brassica ol
41	86	3.6	1627	1 ADPI_MYCPN	P11311 mycoplasma
42	85.5	3.6	742	1 CD44_HUMAN	P16070 h cd44 anti
43	85.5	3.6	790	1 SEIL_MOUSE	O9z2q6 mus musculu
44	85	3.6	417	1 DHMH_PARDE	P29894 paracoccus
45	85	3.6	517	1 EAS_DROME	P54352 drosophila

ALIGNMENTS

RESULT 1					
PAL_KLEPN	ID	PAL_KLEPN	STANDARD;	PRT;	286 AA.
AC	P37446;				
DT	01-OCT-1994 (Rel. 30, Created)				
DT	01-OCT-1994 (Rel. 30, Last sequence update)				
DT	01-NOV-1997 (Rel. 35, Last annotation update)				
DE	PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT				
DE	PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-				
DE	ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).				
GN	PIDA.				
OS	Klebsiella pneumoniae.				
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;				
OC	Klebsiella.				
OX	NCBI_TaxID=573;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=94131966; Pubmed=8300539;				
RA	Brok R.G.P.M.; Brinkman E.; van Bostel R.; Bekkers A.C.A.P.,				
RA	Verheij H.M.; Tommassen J.;				
RT	"Molecular characterization of enterobacterial pida genes encoding				
RT	outer membrane phospholipase A.,"				
RL	J. Bacteriol. 176:861-870(1994).				
CC	-1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE				
CC	A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.				
CC	-1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O -> 1-ACYLGLYCERO-				
CC	PHOSPHOCHOLINE + A FATTY ACID ANION.				
CC	-1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O -> 2-ACYLGLYCERO-				
CC	PHOSPHOCHOLINE + A FATTY ACID ANION.				
CC	-1- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.				
CC	-1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES				
CC	LOCATED THERE.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; X76901; CAA54223.1; ..				
DR	PIR; B36971; B36971.				
DR	PIR; S40129; S40129.				
DR	InterPro; IPR003187; P1A1.				
DR	Pfam; PF02253; P1A1; 1.				
DR	Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium.				
FT	SIGNAL 1 20 BY SIMILARITY.				
FT	CHAIN 21 286 PHOSPHOLIPASE A1.				
FT	ACT_SITE 161 161 BY SIMILARITY.				
SQ	SEQUENCE 286 AA; 32544 MW; 3E39F863085108A3 CRC64;				

Query Match 15.8%; Score 373; DB 1; Length 286;
Best Local Similarity 42.1%; Pred. No. 8.1e-23;
Matches 90; Conservative 27; Mismatches 87; Indels 10; Gaps 5;

RL Nature 409:529-533(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
R Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [6]
RP SEQUENCE OF 30-289 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=85003590; PubMed=6383820;
RA de Geus P., Verheij H.M., Riegan N.H., Hoekstra W.P.M., de Haas G.H.;
RT "The pro- and mature forms of the E. coli K-12 outer membrane
phospholipase A are identical.";
RL EMBO J. 3:1799-1802(1984).
RN [7]
RP SEQUENCE OF 174-289 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=87115164; PubMed=3027506;
RA Irino N., Nakayama K., Nakayama H.;
RT "The recQ gene of Escherichia coli K12: primary structure and
evidence for SOS regulation.";
RL Mol. Gen. Genet. 205:298-304(1986).
RN [8]
RP MUTAGENESIS OF SER-172.
RX MEDLINE=94131966; PubMed=8300539;
RA Brok R.G.P.M., Brinkman E., van Bostel R., Bekkers A.C.A.P.,
RA Verheij H.M., Tommassen J.;
RT "Molecular characterization of enterobacterial plda genes encoding
outer membrane phospholipase A.";
RL J. Bacteriol. 176:861-870(1994).
RN [9]
RP ACTIVE SITE SER-164, AND PARTIAL SEQUENCE.
RX MEDLINE=91249806; PubMed=2040286;
RA Horrevorts A.J.G., Verheij H.M., de Haas G.H.;
RT "Inactivation of Escherichia coli outer-membrane phospholipase A by
the affinity label hexadecanesulfonyl fluoride. Evidence for an
active-site serine.";
RL Eur. J. Biochem. 198:247-253(1991).
CC A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.
CC REQUIRED FOR EFFICIENT SECRETION OF BACTERIOCIINS, SEEMS TO BE
CC DORMANT IN NORMAL GROWING CELLS.
CC -!- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -!- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -!- COPACITOR: REQUIRES CALCIUM IONS FOR ACTIVITY.
CC -!- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES
CC LOCATED THERE.
CC -!- INDUCTION: BY MEMBRANE DAMAGING, FOR EXAMPLE, BY PHAGE-INDUCED
CC LYSIS OR TEMPERATURE SHOCK.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02143; CAA26081.1; -;
DR EMBL; M87049; AAA67617.1; -;
DR EMBL; AE000458; AAC76824.1; -;
DR EMBL; AE005613; AAC59017.1; -;
DR EMBL; AP002567; BAB38174.1; -;
DR EMBL; M30198; AAA24516.1; -;
PIR: A00771; PSECA.

DR PIR; A22133; PSECA1.
DR PIR; S30711; S30711.
DR ECoGene; EG10738; plda.
DR InterPro; IPR003187; PLA1.
DR Pfam; PF02253; PLA1; 1.
KW Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium;
KW Complete proteome.
FT SIGNAL 1 20
FT CHAIN 21 289
FT ACT_SITE 164 164
FT MUTAGEN 172 172
FT CONFLICT 14 15
FT CONFLICT 30 33
FT CONFLICT 30 33
SQ SEQUENCE 289 AA; 33163 MW; A688AD32AA60F218 CRC64;

Query Match 15.18; Score 357; DB 1; Length 289;
Best Local Similarity 39.38; Pred. No. 1.6e-21;
Matches 84; Conservative 29; Mismatches 91; Indels 10; Gaps 4;

QY 226 RAPELKFOVSVKVKAEDLWGTDSLDWFGYTQOSHQWIFNGKNSRPFVRVHDYQEIFLTQ 285
DB 83 RKDEVKFOLSLAFPLWRGILGNPSVLGASYTKQKSWQLSNSESSFFRETNYEPQLFLGF 142
QY 286 PV-YSDLPWDGKVRMIGMGAHVHSHNGESAKLSRSNNRAYLMAGMEKMLTVMPIWGRIF 344
DB 143 ATDYRFAGW--TLRDVEMGYNHDSNGRSDPTSRSNRLYTRLMAENGWLVKPKWYV- 199
QY 345 KEGSGSQPDNDPDLIDYGYGDVRELYOLEKNSITGIVRYNPRSGKALQLDYVYPLGK 404
DB 200 -----GNTDDNDPDTKMGYQLKTYGHL-GDAVLSAKGQVNMWNTGYGAELGLSPYTK 253
QY 405 GISGVFQIFQGYGQSLIDYNEHTEATSGVGLMLND 438
DB 254 HVRLYQVYSGESLIDYNEHTEATSGVGLMLND 287

RESULT 4
PAL_PROVU
ID PAL_PROVU STANDARD; PRT; 289 AA.
AC P37447;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT
DE PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-
DE ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).
GN PLDA.
OS Proteus vulgaris.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Proteus.
OX NCBI_TaxID=585;
RP SEQUENCE FROM N.A.
RX MEDLINE=94131966; PubMed=8300539;
RA Brok R.G.P.M., Brinkman E., van Bostel R., Bekkers A.C.A.P.,
RA Verheij H.M., Tommassen J.;
RT "Molecular characterization of enterobacterial plda genes encoding
outer membrane phospholipase A.";
RL J. Bacteriol. 176:861-870(1994).
CC -!- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE
CC A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.
CC -!- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION
CC -!- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -!- COPACITOR: REQUIRES CALCIUM IONS FOR ACTIVITY.
CC -!- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES
CC LOCATED THERE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its


```
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90105403; PubMed=2690949;
RA Meyer S.L., Kvalnes-Krick K.L., Schramm V.L.;
RT "Characterization of AMD, the AMP deaminase gene in yeast. Production
of amd strain, cloning, nucleotide sequence, and properties of the
protein.";
RL Biochemistry 28:8734-8743(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Badcock K., Churcher C., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY
    METABOLISM.
CC -1- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).
CC -1- PATHWAY: PURINE NUCLEOTIDE CYCLE.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: BELONGS TO THE ADENOSINE AND AMP DEAMINASES FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL; M30449; AAA34420.1; -
DR EMBL; A46659; CAA86620.1; -
DR PIR; A33365; A33365.
DR SGD; S0004498; AMD1.
DR InterPro: IPR001365; A_deaminase.
DR Pfam; PF00962; A_deaminase; 1.
DR PROSITE; PS00485; A_DEAMINASE; 1.
KW Hydrolyase; Nucleotide metabolism.
FT ACT_SITE 422 422 POTENTIAL.
FT ACT_SITE 631 631 POTENTIAL.
FT ACT_SITE 707 707 POTENTIAL.
FT ACT_SITE 708 708 POTENTIAL.
FT CONFLICT 568 568 F -> C (IN REF. 1).
SQ SEQUENCE 810 AA; 93301 MW; 7A6DCB43B9B45C93 CRC64;
```

```
Query Match 4.28; Score 98.5; DB 1; Length 810;
Best Local Similarity 21.38; Pred. No. 3.9; Indels 103; Gaps 15;
Matches 74; Conservative 39; Mismatches 132;
```

```
Qy 46 DNELPIDVQSATQASDTAN-----PLDEHEPELYTTALENTML 86
| : : : : : : : : : : : : : : : : : : : : : : : :
Db 48 DETPPEQDSHESLAADSRANFSYHENQOLLENGTKQLALDEHDS--HSAILEQPSHS 105
| : : : : : : : : : : : : : : : : : : : : : : : :
Qy 87 INCSALNODIMRLACYDTLVH-----GETPAVITKRSIRLDETITWOTIKGKQVVIQET 141
| : : : : : : : : : : : : : : : : : : : : : : : :
Db 106 TNCSSSNTAAMKNG-HDSADHASQSGKPRTLASAQHILPETL-KSFAGAPVNVKQVR 163
| : : : : : : : : : : : : : : : : : : : : : : : :
Qy 142 TDPIFLAG-----NEKGLTKKDAKQLEYAAKQFTPLSLSDLDLRNNTPL 186
| : : : : : : : : : : : : : : : : : : : : : : : :
Db 164 TSAYKMGMLADDAASQOFLDPPSELIDLYSKVAECRLRAKYQTIISVQNDQDNPKNKG 223
| : : : : : : : : : : : : : : : : : : : : : : : :
Qy 187 W--SSRPHNPMY-----VLPFIHMGKPNRSPNTPSHEAKQFTNEFRAPELKFQSVKV 238
| : : : : : : : : : : : : : : : : : : : : : : : :
Db 224 WVYPPPKPSYNSDTKTVPV-----TNKPDAEVDFTKCE-----260
| : : : : : : : : : : : : : : : : : : : : : : : :
Qy 239 KAAEDLMGTDSDLNFGYQQSHWQIFNGKNSRPRFRVHDYQPEIFLTQPVYSDLPMDGKVR 298
| : : : : : : : : : : : : : : : : : : : : : : : :
Db 261 -----IPGEDPDWEFTLNDDDSYVY-----HRSKTDIELIAQIPTLRDYLDLE-----K 305
| : : : : : : : : : : : : : : : : : : : : : : : :
Qy 299 MIGGVAHVSNGESAK-----LSRSWNRAYLMAGNEWKNLTVMPR 338
| : : : : : : : : : : : : : : : : : : : : : : : :
Db 306 MISIS-----SDGPAKSFAYRRLQYLEARWNLVYLL--NEYQETSVSKR 347
| : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 7
YJHO_YEAST
ID YJHO_YEAST STANDARD; PRT; 888 AA.
AC P40361;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE HYPOTHEICAL 104.3 KDA PROTEIN IN SMC3-MRPL8 INTERGENIC REGION.
GN YJL070C OR J1095 OR HRD888.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=95282514; PubMed=7762302;
RA vandenbol M., Durand P., Dion C., Portetelle D., Hilger F.;
RT "Sequence of a 17.1 kb DNA fragment from chromosome X of
Saccharomyces cerevisiae includes the mitochondrial ribosomal protein
L8.";
RL Yeast 11:57-60(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Sor F.J.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE ADENOSINE AND AMP DEAMINASES FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL; Z34288; CAA84052.1; -
DR EMBL; Z49345; CAA89362.1; -
DR EMBL; Z49344; CAA89361.1; -
DR EMBL; X88851; CAA61309.1; -
DR PIR; S47120; S47120.
DR SGD; S0003606; YJL070C.
DR InterPro: IPR001365; A_deaminase.
DR Pfam; PF00962; A_deaminase; 1.
KW Hypothetical protein; Hydrolase.
SQ SEQUENCE 888 AA; 104263 MW; DB31A8086224114D CRC64;
```

```
Query Match 4.18; Score 96; DB 1; Length 888;
Best Local Similarity 22.58; Pred. No. 7.1; Indels 126; Gaps 15;
Matches 67; Conservative 35; Mismatches 126;
```

```
Qy 128 QTIKKGPOVIYQE-----TTDPIFLMGNEKGLTKKDA-----KQLEYAAKQFTPLSL 176
| : : : : : : : : : : : : : : : : : : : : : : : :
Db 2 QAVERRPSLLFDEYQNSVTKPNETKNKEARVLSENDGVSPSVLKQKEISVDDMDMISLP 61
| : : : : : : : : : : : : : : : : : : : : : : : :
Qy 177 FDLDRN---NTPLW---SSRPHNPMYVLPFIHMGKPNR-----SPNTFSH-----EAKQF 220
| : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 TEFDRQWVLGSPMFFDLEDEENKIDPLPSVSHHYNGESDSFVSSTYPSNLKGTGETKDL 121
| : : : : : : : : : : : : : : : : : : : : : : : :
Qy 221 TPNFRAPELAFQVSVKVAEDLWG-----TDSLWFGYQ---QSHWQIFNGKNSRPR 273
| : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 FINPF---ELVSQMKRYIAASKQDQISNINKNDEKWEFLYPKPLPKFRFEDDK-----R 173
| : : : : : : : : : : : : : : : : : : : : : : : :
Qy 274 VHDYQPEIFLQPVYSDLPMDGKVRMIGMGAHVHNSGESAKLSRSWNRAYLMAGNEWKNL 333
| : : : : : : : : : : : : : : : : : : : : : : : :
Db 174 FQD-----PDSLDLDDGDSGTGTGAATPHRHGYYPPSYFTDHYYYTKSLGKLGKN 224
| : : : : : : : : : : : : : : : : : : : : : : : :
Qy 334 TVMPRIWGRIFKEGSGSDPDNDPILDYGYGVDVRFYQLENKNSISGTVRNPNSGK 391
| : : : : : : : : : : : : : : : : : : : : : : : :
Db 225 IKVP-YTGEY-----DLEDY-----KKOYIVHLSNQENTONPL--SPYSSK 263
| : : : : : : : : : : : : : : : : : : : : : : : :
```



```

Db 237 CTKNVYSABSHGTSFASDSESPGNSLKPGLPMQGIKSEPPSEETEEESTITLUSGDE 296
QY 125 TIWQTIKGRQVIYQETTDPI-----FLMNEKGMLTK 157
Db 297 T---DIKDRPGDVEMDRKQSPARTSTRTGAACLDLRSVSPSCRLSLFGITKGV--- 350
QY 158 KDAKQLEYAAKQFTPLSLF-----DLDRNNTPLWSS--RPHNPMYVLPFLMHG 204
Db 351 -ESTGLPSTSQ--PLVRSSACPFNKGISQGLKTDYTPLAGNYGPHVQKDVSNFAMG 407
QY 205 KNRSP-----NTPSHEAKQFTNFEAPELK-----FQVSVKVKAEDLWGTDS 250
Db 408 SPLRGPETLCEFFSSPCSGARFLATEHQEPGLMGDMYNOVRPQIK-CEOSYGTNSS 466
QY 251 LMFQYQQSHQIFNGKNSRPRVHDYQPEIFLTQPV--YSDLPMWKGKVRMIGMGAVHHS 308
Db 467 DESG-----SFSADSESCPVDRGQEVKLPFPVDQITDLPRNDFQMMIKM---HKL 515
QY 309 NGESAKL-----SRSNR 321
Db 516 TSQLEFIHDIRRSKRN 533

RESULT 11
SRK6_BRAOL STANDARD; PRT; 849 AA.
ID SRK6_BRAOL STANDARD; PRT; 849 AA.
AC Q09092;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE PUTATIVE SERINE/THREONINE KINASE RECEPTOR PRECURSOR (EC 2.7.1.37)
DE (S-RECEPTOR KINASE) (SRK).
GN SRK6.
OS Brassica oleracea (Cauliflower).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3712;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. S656; TISSUE=Stigma;
RX MEDLINE=92020942; PubMed=1681543;
RA Stein J.C., Howlett B., Boyes D.C., Nasrallah M.E.;
RT "Molecular cloning of a putative receptor protein kinase gene encoded
RT at the self-incompatibility locus of Brassica oleracea.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:8816-8820(1991).
CC 1- FUNCTION: INVOLVED IN SPOPHYTIC SELF-INCOMPATIBILITY SYSTEM
CC (THE INABILITY OF FLOWERING PLANTS TO ACHIEVE SELF-
CC FERTILIZATION), PROBABLY ACTING IN COMBINATION WITH S-LOCUS-
CC SPECIFIC GLYCOPROTEINS. INTERACTION WITH A LIGAND IN THE
CC EXTRACELLULAR DOMAIN TRIGGERS THE PROTEIN KINASE ACTIVITY OF THE
CC CYTOPLASMIC DOMAIN.
CC 1- CATALYTIC ACTIVITY: ATP + A PROTEIN -> ADP + A PHOSPHOPROTEIN.
CC 1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC 1- TISSUE SPECIFICITY: PREDOMINANTLY IN THE PISTIL AND ANTHR.
CC 1- POLYMORPHISM: THERE ARE A NUMBER OF DIFFERENT S ALLELES IN
CC B.OLERACEA, POSSIBLY PROVIDING THE RECOGNITION SPECIFICITY.
CC 1- SIMILARITY: THE EXTRACELLULAR DOMAIN IS SIMILAR TO S-LOCUS
CC GLYCOPROTEINS OF BRASSICA, WHILE THE INTRACELLULAR DOMAIN IS
CC A SER/THR-PROTEIN KINASE RELATED TO RAF KINASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: M76647; AAA33000.1; ALT_TERM.
CC HSP: P11362; IFGI.
CC InterPro: IPR001480; B_lectin.
CC InterPro: IPR000719; Euk_pkinase.

```

```

DR InterPro: IPR003609; Pan_app.
DR InterPro: IPR002290; Ser_thr_kin_actsite.
DR InterPro: IPR000858; Slocus_glycop.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF00954; Slocus_glycop; 1.
DR SMART: SM00108; B_lectin; 1.
DR SMART: SM00473; PAN_AP; 1.
DR SMART: SM00221; STYK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
KW Transferrase; Serine/threonine-protein kinase; Signal; ATP-binding;
KW Transmembrane; Receptor; Glycoprotein; Self-incompatibility.
FT SIGNAL 1 32
FT CHAIN 33 849
FT PUTATIVE SERINE/THREONINE KINASE
FT RECEPTOR.
FT EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 447 466
FT DOMAIN 447 466
FT DOMAIN 467 849
FT DOMAIN 528 779
FT NP_BIND 534 542
FT BINDING 556 556
FT ACT_SITE 653 653
FT CARBOHYD 47 47
FT CARBOHYD 120 120
FT CARBOHYD 196 196
FT CARBOHYD 260 260
FT CARBOHYD 314 314
FT CARBOHYD 389 389
FT CARBOHYD 442 442
SQ SEQUENCE 849 AA; 97231 MW; 7E156059EDDF4370 CRC64;

Query Match 3.9%; Score 92.5; DB 1; Length 849;
Best Local Similarity 21.6%; Pred. No. 13; Mismatches 107; Indels 121; Gaps 24;
Matches 77; Conservative 51;

QY 105 LVHGETPAVTKRSIRLDETITWQTIKGRQVIYQETTDPIFLMG----NEK---GMLTK 157
Db 24 LIHPALSIYINT-----LSTESTLTISSNKTIV---SPGSEFVGVFERTNSRWYLGWYK 75
QY 158 K-DAKQLEYAAKQFTPLSLF-----DLDRNNTPLWSSRPHNPMYVLPFLMHGK 205
Db 76 KYSDRTYVNVANRDNPLSNAIGTLKISGNLVLDDHSNKPVWNTN-----LTRGN 125
QY 206 PNRSP-----NTPSHEAKQFT-----PNEFRAPELKFQVSVKVKAEDLW 245
Db 126 -ERSPVVAELLANGNVFMRDSSNNDASEYLWQSFDPYTDLLPEMKLGYNLKT----- 177
QY 246 GTDSDLWFGYTOQSHWQIFNGKNSRPRVHDYQ-----PEIFLTQ---PVYSDLPWDGK 296
Db 178 GLNRF-----TSWRSDDDPSSGNF---SYKLETQSLPEFVLSRENPFMRHSGPWNG- 226
QY 297 VRMIGMGAVH-----HNGESAKLSRSWNRAY-----LMAGMEKNTLVMP--RI 339
Db 227 IRFSGIPEDQKLSYMYNFENNEEVAYTFRTMNSNFSRLTSLISGYQLRTWYPSIRI 286
QY 340 WGRIFREGSGSOPDDNPDILDY-----YGVGVRELYQLENKSNISGTVR-VNPRS 389
Db 287 WNRFW-----SSPVD-PQCDTYIMCGPYACDV-----NTSPVCNCIQGFNPRN 329

RESULT 12
YL70_ARCFU STANDARD; PRT; 324 AA.
ID YL70_ARCFU STANDARD; PRT; 324 AA.
AC O28112;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN AF2170.
GN AF2170
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;

```

```
OC Archaeoglobus.
OX NCBI_TaxID=2234;
FN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klench H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Quake A.R., Graham D.E., Kyriades N.C.,
RA Fleischmann R.D., Kachakianush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -1- SIMILARITY: CONTAINS 2 KELCH REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000955; AAB89093.1; -
DR TIGR; AF2170; -
DR InterPro; IPR001798; Kelch.
DR Pfam; PF01344; Kelch; 2.
KW Hypothetical protein; Repeat; Complete proteome.
FT REPEAT 229 276 KELCH 1.
FT REPEAT 277 323 KELCH 2.
SQ SEQUENCE 324 AA; 36025 MW; C445388CFEB96B45 CRC64;

Query Match 3.9%; Score 92; DB 1; Length 324;
Best Local Similarity 20.1%; Pred. No. 3.7;
Matches 54; Conservative 44; Mismatches 72; Indels 98; Gaps 18;

QY 224 EFRAPELKFQSVKVKAAED-----LWGTDSLDFGTYQSHWQIF-NGKNSRPRVR 275
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 63 EFHPNP-RLEI---LLSSEDFSQREAAVDGEILLIFGGT-----VFENGKYSPTDQIL 112
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 276 DYQPEIFLTPVYSDLP-----W-DGKVRMI-----GMCAYVHNGESAKLSRS 318
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 113 SFNPKLERLRVNASLPHPTSDVAAVWGDGRVYIFLNNSECEVYAFYPSNESFAKLDSV 172
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 319 -----W--NRAYL-----WAGWEKNLTVMPRIWGRIFKESGSGQ 351
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 173 CPHEPGGCVHVVWYGGKAYFFGEGVASFDPMGGEKFW--IAFTDRVWVRAATVADG-- 228
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 352 PDONPDILYYGYGDVFLVLENKSNISGT---VYNPGRSGKALQLDVYVPLGKG-- 405
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 229 -----YIFAIGSSGIAETKDEIIRFNFTGE-LCEMTKLPVARGQA 270
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 406 --TSG-YEQIF-----QGYQSLIDYNH 425
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 271 VAVGGEYIYFGYTKDGYANEILLRVDY 298
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
RI14_HUMAN STANDARD; PRT; 1158 AA.
AC P48552;
DT 01-FEB-1996 (Rel. 33, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE NUCLEAR FACTOR R1P140 (NUCLEAR RECEPTOR INTERACTING PROTEIN 1).
```


Qy 421 -----IDYNH 425
: : : :
Db 417 EKASREQQLDYSH 429

Search completed: November 30, 2001, 14:27:00
Job time: 547 sec

This Page Blank (uspto)


```
QY 141 TTDPIFLMGNEKGLTKKDAKQLEYAAKQFTPLSLSFDLDRNN-TPLWSSRRPHNPMYVLP 199
Db 93 GGDAL-----PADSAGETADIVTPLSLMYDLKNDLRLGLLVREHNPYMLP 139

QY 200 IFMHGKPNRSPNTPSH-EAKQFTPNFRAPDLKQVSVKVAEDLWGTDSDLWFGYTQ 258
Db 140 LWTNNSPNYAPGSPTRGTTTVOEKFGQOKRAETKLOVSFKIAEDLTKTRADLWFGYTQ 199

QY 259 SHWQIFN-GKNSRPRVHDYQPEIFLTQPVYSDLPWDGKVRMIGMGAVHHSNGESAKLSR 317
Db 200 SDWQIYNQGRKSAPFRNTDYKPEIFLTQPVKADLPFGGRMLRMIGAGFVHOSNGOSRPSR 259

QY 318 SWRAYLMAGMEWKNLTVMPRIMGRIFKEGSGSQPDNDPILDYGYGDRFLYQLENKS 377
Db 260 SWNRIVAMAGMEWGNKLTVIPRVVWVAFDQ-SGDK-NDNPDIAVMGYGDKVQLYRLNDRQ 317

QY 378 NISGTVRYNPRSGKALQLDYVYPLKGISGYFOIFOGYQOSLIDYNHEATSFVGGLMLN 437
Db 318 NVYSVLRYNPKTYGAEAAATTPKIKGLKGVVRGFGHGYGESLIDYNHKQNGIGIGLMFN 377

QY 438 DWMGL 442
Db 378 DLGDI 382

RESULT 2
Q9JUT21
ID Q9JUT21 PRELIMINARY; PRT; 409 AA.
AC Q9JUT21
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE PUTATIVE PHOPHOLIPASE.
GN NWA2021.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
Whitehead S., Pratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis 22491."
RL Nature 404:502-506(2000).
DR EMBL: AL162757; CAB85240.1; -.
DR InterPro: IPR003187; PLAI.
DR Pfam: PF02253; PLAI; 1.
KW Complete proteome.
SQ SEQUENCE 409 AA; 45862 MW; CD65858064D01AA1 CRC64;
```

```
Query Match 34.5%; Score 814; DB 2; Length 409;
Best Local Similarity 44.4%; Pred. No. 1.2e-62;
Matches 162; Conservative 68; Mismatches 111; Indels 24; Gaps 8;

QY 87 INCSALNODIMRLACDYTLVHGTEPAVI-----KTKRSIRLDEFINQTI-KGKQVYIQE 140
Db 60 LQCAALTDNVTRLACDYRIFAQALPSAGQEGESKAVLNLTETVRSRLSDKGEAVIVVEK 119

QY 141 TTDPIFLMGNEKGLTKKDAKQLEYAAKQFTPLSLSFDLDRNN-TPLWSSRRPHNPMYVLP 199
Db 120 GGDAL-----PADSAGETADIVTPLSLMYDLKNDLRLGLLVREHNPYMLP 166

QY 200 IFMHGKPNRSPNTPSH-EAKQFTPNFRAPDLKQVSVKVAEDLWGTDSDLWFGYTQ 258
Db 167 LWTNNSPNYAPGSPTRGTTTVOEKFGQOKRAETKLOVSFKIAEDLTKTRADLWFGYTQ 226
```

```
QY 259 SHWQIFN-GKNSRPRVHDYQPEIFLTQPVYSDLPWDGKVRMIGMGAVHHSNGESAKLSR 317
Db 227 SDWQIYNQGRKSAPFRNTDYKPEIFLTQPVKADLPFGGRMLRMIGAGFVHOSNGOSRPSR 286

QY 318 SWRAYLMAGMEWKNLTVMPRIMGRIFKEGSGSQPDNDPILDYGYGDRFLYQLENKS 377
Db 287 SWNRIVAMAGMEWGNKLTVIPRVVWVAFDQ-SGDK-NDNPDIAVMGYGDKVQLYRLNDRQ 344

QY 378 NISGTVRYNPRSGKALQLDYVYPLKGISGYFOIFOGYQOSLIDYNHEATSFVGGLMLN 437
Db 345 NVYSVLRYNPKTYGAEAAATTPKIKGLKGVVRGFGHGYGESLIDYNHKQNGIGIGLMFN 404

QY 438 DWMGL 442
Db 405 DLGDI 409

RESULT 3
Q9CL22
ID Q9CL22 PRELIMINARY; PRT; 306 AA.
AC Q9CL22;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL PROTEIN PM1426.
GN PM1426.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RT May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
"Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL: AE006179; AAK03510.1; -.
DR InterPro: IPR003187; PLAI.
DR Pfam: PF02253; PLAI; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 306 AA; 35580 MW; EAF3DE8C1C22B26E CRC64;

Query Match 15.8%; Score 372; DB 2; Length 306;
Best Local Similarity 39.0%; Pred. No. 2.5e-24;
Matches 83; Conservative 40; Mismatches 82; Indels 8; Gaps 4;

QY 229 ELKFQSVKVKAAEDLWGTDSDLWFGYTQOSHWOIFENGKNSRPRVHDYQPEIFLTQPVY 288
Db 97 EIRFKISLALPLWRGILGNNSVLAASYTKSWFOLSNDSSPPRETNYEPQLFLANKTQ 156

QY 289 SDLPWDGKVRMIGMGAVHHSNG--ESAKLSRMNRYLMAGMEWKNLTVMPRIMGRIFKE 346
Db 157 YSLPFGTWLQDVETGINHQSGNRDDEAKLSRNRLVVRASAIKQNTVTEIKPMWRIPEK 216

QY 347 GSGSQPDNDPILDYGYGDRVP-LYOLENKSNTSGTVRYNPRSGKALQLDYVYPLK 405
Db 217 ---AKNDNDPDITKYRGHFDVALGYHYHDFKLSG--HYNPTSNKGGLEASYPITKN 271

QY 406 ISGVFOIFOGYQOSLIDYNHEATSFVGGLMLND 438
Db 272 IRFTQYNYGSLIDYQQRIRIGIGISLNN 304

RESULT 4
Q9PMU8
ID Q9PMU8 PRELIMINARY; PRT; 329 AA.
AC Q9PMU8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
```


QY 345 KEGSGQDDNDPILDYGGVDVRFYQLENKSNISCTVRYNPRSGKALQLDYVYPLGK 404
 DB 200 -----GSTDDNDPITKMGYQKICGYHL-GEAVLSAKGQYNWNTGYGAEVGLSYPVTK 253
 QY 405 GISGYQIFQGYGQSLIDYNHEATSFVGVLMLND 438
 DB 254 HVRLYQVYSGYGESLIDYNFNQTRVGVGVLMLND 287
 RESULT 7
 ID Q924N8 PRELIMINARY; PRT; 292 AA.
 AC Q924N8;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE OUTER MEMBRANE PHOSPHOLIPASE A PRECURSOR (EC 3.1.1.32).
 GN PLDA.
 OS Enterobacter agglomerans.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Pantoea.
 OX NCBI_TaxID=549;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94131966; PubMed=8300539;
 RA Brok R.G., Brinkman E., van Bostel R., Bekkers A.C., Verheij H.M.,
 RA Tommassen J.;
 RT "Molecular characterization of enterobacterial plda genes encoding
 RT outer membrane phospholipase A";
 RL J. Bacteriol. 176:861-870(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99120145; PubMed=9921577;
 RA Brok R.G., Boots A.P., Dekker N., Verheij H.M., Tommassen J.;
 RT "Sequence comparison of outer membrane phospholipases A: implications
 RT for structure and for the catalytic mechanism";
 RL Res. Microbiol. 149:703-710(1998).
 DR EMBL; AF034414; AAD03498.1; -;
 DR InterPro; IPR003187; PLAI.
 DR Pfam; PF02253; PLAI; 1.
 DR PRINTS; PR01486; PHPLIPASEA1.
 KW Signal; Hydrolase.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 292 OUTER MEMBRANE PHOSPHOLIPASE A.
 FT SEQUENCE 292 AA; 33719 MW; B75516D093B2BEEA CRC64;
 Query Match 15.1%; Score 356.5; DB 2; Length 292;
 Best Local Similarity 40.1%; Pred. No. 5.2e-23;
 Matches 85; Conservative 31; Mismatches 87; Indels 9; Gaps 5;
 QY 229 ELKFOVSKVKAEDLWGTDSLWFGYQSHQIFNGKNSRPRVHDYQPEIFLTQPV- 287
 DB 86 EVKFLSLGFPPIWRGIAGDNLGASYSYQSWQASNDSPFRETNYEPQIFLAWATD 145
 QY 288 YSDLPWDGKVRMIGMGAVHNSGESAKLSRWNRAVLMAGMKNLTVMPIRIGRIFKEG 347
 DB 146 YELAGW--TFREVEFGYHNSGKADPTSRSDRYVTRLMAGQGNLEIDLPWRIPESD 203
 QY 348 SGSGQDDNDPILDYGGVDVRFYQLENKSNISCTVRYNPRSGKALQLDYVYPLGKI 406
 DB 204 S---KDDNPDKYMGYRLKVGALGESVFSIDG--RYNWNNTGYGGAEGWSYPTIKHV 258
 QY 407 SGYQIFQGYGQSLIDYNHEATSFVGVLMLND 438
 DB 259 RFYTOVPSGYGESIDYNFRQTRVGVGVLMLND 290
 RESULT 8
 ID Q32349 PRELIMINARY; PRT; 297 AA.
 AC Q32349;

DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE PHOSPHOLIPASE A.
 GN PLDA.
 OS Campylobacter coli.
 OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
 OC Campylobacter.
 OX NCBI_TaxID=195;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UAS85;
 RX MEDLINE=97230284; PubMed=9119448;
 RA Grant K.A., Belandria I., Dekker N., Richardson P.T., Park S.F.;
 RT "Molecular characterization of plda, the structural gene for a
 RT phospholipase A from Campylobacter coli, and its contribution to cell-
 RT associated hemolysis";
 RL Infect. Immun. 65:1172-1180(1997).
 DR EMBL; Y11031; CAA71915.1; -;
 DR InterPro; IPR003187; PLAI.
 DR Pfam; PF02253; PLAI; 1.
 DR PRINTS; PR01486; PHPLIPASEA1.
 SQ SEQUENCE 297 AA; 34998 MW; 04B54A7BCA3764CE CRC64;
 Query Match 14.5%; Score 342.5; DB 2; Length 297;
 Best Local Similarity 32.4%; Pred. No. 8.8e-22;
 Matches 97; Conservative 37; Mismatches 108; Indels 57; Gaps 11;
 QY 178 DLDRNNTPLSSRPHNPMYVLPFMHGKPNRSPNT-----P 213
 DB 16 DLKNNASLLSRKHETQ-----NTQKTPSKEDFSRIALANYLGENSESFPNPLGIS 66
 QY 214 SHEAKOFTP-----NEFRAPELKFOYSVKAAEDLWGTDSLWFGYQSHQIF 264
 DB 67 SYKMNVFLPAYSGSLGGENRKTETMKFQLSKRLEFEDLLGLGEKYVGTQTSWQ-- 124
 QY 265 NGKNSRPRVHDYQPEIFLTQPV--YSDLPWDGKVRMIGMGAVHNSGESAK--LSRSWNR 321
 DB 125 NYKHSSPFRETNYQPEFFVDIPLHFEDYKFLNNLR---VGLHESNGKGDENLESRSWNR 181
 QY 322 AYLMAGMEWKNLTVMPIRIGRIFKEGSGSQDDNDPILDYGGVDVRFYQLENKSNISG 381
 DB 182 IYASSVELYQRFVLPRIWYRI---PENSDDDNPEITHYMGNPDIN-IGSLGNDYFINL 237
 QY 382 TVRYNP--RSGKALQLDYVYPL--GKISGYQIFQGYGQSLIDYNHEATSFVGVLMLN 437
 DB 238 MLRNNLDFHDKGAVQVDIGYDIFDNGIYWLQYFNGYGGSLIDYNKRLQLRSLTAFLIS 296
 RESULT 9
 ID Q25241 PRELIMINARY; PRT; 355 AA.
 AC Q25241;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE PHOSPHOLIPASE A1 PRECURSOR (DR-PHOSPHOLIPASE A).
 GN HP0499
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter.
 OX NCBI_TaxID=210;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=26695 / ATCC 700392;
 RX MEDLINE=97394467; PubMed=9252185;
 RA Tomb J.-F., White O., Kerlavage A.R., Klenk H.-P., Gill S., Dougherty B.A.,
 RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Peterson S.,
 RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
 RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
 RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
 RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,

6	TKQLISSILILVL--LLSLFPTALAAEGNTREDNFKHLGNDNVKRRSEAGALQLEVDGQ	64
66	NPL-DEHEPELY-----TTALENTMLINCSAL-----NQDIMRLACY-DTLVHGTEP	111
65	MTLVDQHGEEKIQLRGMTSHGLQWFPETLNDNAVYKALANDWESNMIRLAMYVGNGYASNP	124
112	AVIKTKRSIRLDETI-----WQT-----IKGKQVIV	138
125	ELIKSRVIKGIDLAIENDMVIVDVHWHAPDRDPVYAGAEDEFDRIALYPNNPHIY	164
139	QETTDP-----ITFLMGNEKGMILTKDKAKOLEYAAKQFTPLSLSFOLDRN-----NTPLW	187
185	ELANEPSSNNGGAGIPNNEEGWNAV-----EYADPIVEMLRDSGNADONIIIVGSPNW	239
188	SSRP-----HNPWVYLPIT--MHGKPNSS--PNTPSHKAQFTPNEFRAPELKFQ	233
240	SQRDLAADNPIDOHHTMYTVHFYTGTHAASTESYPETPNSRGVNVMSNTRYA-----LE	295
234	VSVKVAEADLWGT-----DSDLWFGYTOQ-----SHWQIFNGKNS-----RP	271
296	NGVAVEATE--WGTISQANGDGGPYDEADWIEFLNENNISWANNSLTN-KNEVSGAFTP	352
272	FRVH-----DYQPELFTQPYVSOLPW---D	294
353	FELGKSNATSLDPGPDQVWVPEELSLSGEYVRAIRIKGVNYP---IDRTKYTKTVLMDFND	409
295	GKVRMIGMGAHVHNGSS-----AKLSRWNRAYLMAGMEKNNLWVMPRIMG	341
410	GTKGFGV-----NGDSPVEDVVIENEAGALKLSGLDASNDYSEGNVYANARLSADGCG	463
342	R 342	
464	K 464	

```

RESULT 15
ID Q9F216
ID ID Q9F216 PRELIMINARY; PRT; 824 AA.
AC Q9F216;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CELLULOASE.
OS Bacillus sp.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1409;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KSM-S237;
RX MEDLINE=121036886; PubMed=119339393;
RA Hakamada Y., Hatada Y., Koike K., Yoshimatsu T., Kawai S.,
RA Kobayashi T., Ito S.;
RT "Deduced amino acid sequence and possible catalytic residues of a
RT thermotstable, alkaline cellulase from an alkaliophilic Bacillus
RT strain.";
RL Biosci. Biotechnol. Biochem. 64:2281-2289(2000).
DR ENBL: AB018420; BAB19360.1; -.
DR InterPro: IPR001547; Glyco_hydro_F5.
DR Pfam: PF00150; cellulase; 1.
DR PROSITE: PS00659; GLYCOSYL-HYDROL_F5; UNKNOWN_1.
SQ SEQUENCE 824 AA; 91564 MW; 65FA940FE1D729B9 CRC64;

```

```

Query Match          4.7%  Score 110.5;  DB 2;  Length 821;
Best Local Similarity 20.0%;  Pred. No. 0.62;
Matches 96;  Conservative 53;  Mismatches 163;  Indels 169;  Gaps 25;
SQ  SEQUENCE      821 AA;  90910 MW;  73D438FEF0B40B5C  CRC64;
Qy  9  TLTSLSCFAILLIAQQAQVPPNPFVDFVRSND---LGDQNELPIDVQSATQASTDTA  65
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

```

Db 66 MTLVDQGEKIQLRGMSHGLQWFPETLNDNAYKALSNDWDSNMRLAMY-----VGENGX 121
QY 110 --TPAVIKTK-----RSIRLDETI--WQT-----IKGKP 134
Db 122 ATNPELIKQVRIDGIELAIEDMIVVDWVHAPGDRDPVYAGAKDFFREIAALYPNP 181
QY 135 QVIYQETDP-----IFLMGNEKMLTKKDAKQLEAAKQFTPLSLSPDLDRN-----N 183
Db 182 HIIYELANEPSSNNGGAGIPNNEEGKAVK-----EYADPIVEMLRKSGNADDNIIIVG 236
QY 184 TPLWSSRP-----HNPWIVLPFPMHGKPNRSPNTGSHAKQFTPNFRAPELK-- 231
Db 237 SPNWSQRPDLAADNPIDDHHTMTVTFYGTGSHAATESYPSE-----TPNSERGNVMSNT 291
QY 232 ---FOVSVKVKAAEDLWGT-----DSDLWFGYTQQ-----SHQJENGKNS-- 269
Db 292 RYALENGVAVFATE--WGTQASGDGPGPYFDEADVWIEFLNENNISWANNSLTN-KNEVS 348
QY 270 ---RPFVR-----HDYQP-EIFLT-----QPV-----YSDLPW-- 293
Db 349 GAFTPFELGKSNATNLDGPDHVAPEELSLSGEYVRARIKGVNYEPIDRTKYTKVLWDF 408
QY 294 -DGKVRMIGCAVHHNGESAKLSRSNRAYLMAGME-----WKNLTVMPRIWGR 342
Db 409 NDGTKQ--GFGVNSDSPNKELIAVDNENNTLKVSGLDVSNVDSDGNFWANARLSANGWGK 466

```

Search completed: November 30, 2001, 14:26:25
Job time: 562 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:16:58 ; Search time 72.04 Seconds
(without alignments)
454.475 Million cell updates/sec

Title: US-09-787-083-4
 Perfect score: 2360
 Sequence: 1 MKVSI STLTL SLILPCFAILA.....YNHEATSRGVGLMNDWMGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

Database :

A_Geneseq_1101.*

1:	/SID58/gcgdata/geneseq/geneseqp/AA1980.DAT.*
2:	/SID58/gcgdata/geneseq/geneseqp/AA1981.DAT.*
3:	/SID58/gcgdata/geneseq/geneseqp/AA1982.DAT.*
4:	/SID58/gcgdata/geneseq/geneseqp/AA1983.DAT.*
5:	/SID58/gcgdata/geneseq/geneseqp/AA1984.DAT.*
6:	/SID58/gcgdata/geneseq/geneseqp/AA1985.DAT.*
7:	/SID58/gcgdata/geneseq/geneseqp/AA1986.DAT.*
8:	/SID58/gcgdata/geneseq/geneseqp/AA1987.DAT.*
9:	/SID58/gcgdata/geneseq/geneseqp/AA1988.DAT.*
10:	/SID58/gcgdata/geneseq/geneseqp/AA1989.DAT.*
11:	/SID58/gcgdata/geneseq/geneseqp/AA1990.DAT.*
12:	/SID58/gcgdata/geneseq/geneseqp/AA1991.DAT.*
13:	/SID58/gcgdata/geneseq/geneseqp/AA1992.DAT.*
14:	/SID58/gcgdata/geneseq/geneseqp/AA1993.DAT.*
15:	/SID58/gcgdata/geneseq/geneseqp/AA1994.DAT.*
16:	/SID58/gcgdata/geneseq/geneseqp/AA1995.DAT.*
17:	/SID58/gcgdata/geneseq/geneseqp/AA1996.DAT.*
18:	/SID58/gcgdata/geneseq/geneseqp/AA1997.DAT.*
19:	/SID58/gcgdata/geneseq/geneseqp/AA1998.DAT.*
20:	/SID58/gcgdata/geneseq/geneseqp/AA1999.DAT.*
21:	/SID58/gcgdata/geneseq/geneseqp/AA2000.DAT.*
22:	/SID58/gcgdata/geneseq/geneseqp/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	2360	100.0	442	21	AAV855269	BASB034 amino acid
2	2339	99.1	442	21	AAV85271	BASB034 amino acid
3	2330	98.7	442	21	AAV85268	BASB034 amino acid
4	2330	98.7	442	21	AAV85270	BASB034 amino acid
5	822	34.8	370	21	AAV75156	Neisseria gonorrhoe
6	815	34.5	370	21	AAV75157	Neisseria meningit
7	815	34.5	370	21	AAV75158	Neisseria meningit
8	815	34.5	370	21	AAV70629	Neisseria meningit
9	812	34.4	375	21	AAV70628	Neisseria meningit
10	246.5	10.4	355	19	AAV98871	H. pylori GHPO 172
11	239.5	10.1	359	19	AAV910360	H. pylori ORF 07ap

12	156.5	6.6	253	18	AAW20760
13	109.5	4.6	824	21	AAB23180
14	109	4.6	800	8	AAW20760
15	109	4.6	822	13	AAW26021
16	107	4.5	157	18	AAW20538
17	106	4.5	798	21	AAB40925
18	106	4.5	798	21	AAW20538
19	106	4.5	798	22	AAW20538
20	105.5	4.5	537	22	AAW31000
21	102.5	4.3	1686	19	AAW70991
22	98.5	4.2	467	20	AAW13378
23	98.5	4.2	467	21	AAW13378
24	98.5	4.2	467	22	AAW13378
25	98.5	4.2	467	22	AAW13378
26	98.5	4.2	1726	18	AAW38756
27	97	4.1	761	20	AAW90084
28	95.5	4.0	502	22	AAW90084
29	95.5	4.0	516	22	AAW90084
30	95	4.0	522	22	AAW90084
31	94	4.0	372	16	AAW95616
32	93.5	4.0	467	19	AAW5057
33	92	3.9	888	22	AAW70751
34	92	3.9	1484	12	AAW11749
35	91.5	3.9	857	13	AAW29814
36	91	3.9	600	22	AAW75092
37	91	3.9	1227	21	AAW81501
38	91	3.9	1474	21	AAW971157
39	91	3.9	1474	22	AAW50673
40	90	3.8	522	22	AAW95513
41	90	3.8	564	21	AAW10457
42	90	3.8	564	21	AAW81952
43	90	3.8	790	22	AAW93045
44	90	3.8	841	20	AAW24318
45	90	3.8	1627	16	AAW67538

ALIGNMENTS

RESULT 1

AAV85269
ID AAV85269 standard; Protein; 442 AA.

AA
AC
AA85269:

AA
DT 29-JUN-2000 (first entry)

DE BASB034 amino acid sequence #2.

Moraxella catarrhalis infection; BAS8034; diagnosis; staging; treatment; vaccine; bacteriostatic; treatment; prevention; otitis media; sinusitis; nosocomial infection; invasive disease; chronic otitis media; hearing loss; antibacterial drug.

OS Moraxella catarrhalis.

WO200015802-A1.

23-MAR-2000.

14-SEP-1999: 99WO-EP06781.

14-SEP-1998: 98GB-0020002.

PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

AA
PI
Ruelle J;

WPI: 2000-271440/23.

DR N-PSDB; AAA10701.

Novel BASB034 polynucleotides and polypeptides from *Moraxella catarrhalis* used to prepare vaccines against bacterial infection

XX Claim 3; Page 67; 106pp; English.
 PS This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 CC strain Mc2908. The invention relates to BASB034 polypeptides from
 CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
 CC polynucleotides and polypeptides may be employed as research reagents and
 CC material for the discovery of treatments and diagnostics for diseases,
 CC particularly human diseases. They are particularly used to diagnose and
 CC treat M. catarrhalis infections. They can be used for diagnosis of
 CC disease, staging of disease, or determining response of an infectious
 CC organism to drugs. The polynucleotides may be used as a source for
 CC hybridization probes, and for screening of genetic mutations, serotype,
 CC organism or strain identification, identification of mutations in BASB034
 CC sequences, and as components of arrays which are useful for diagnostic
 CC and prognostic purposes. The polypeptides can be used to produce
 CC antibodies. The polypeptides can also be used in vaccine formulations,
 CC and to identify agonists and antagonists. The polypeptides, antibodies,
 CC agonists and antagonists (which are bacteriostatic) are used for the
 CC treatment and prevention of diseases such as otitis media in infants and
 CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
 CC invasive diseases, and chronic otitis media with hearing loss. The
 CC polypeptides, agonists and antagonists are also used for screening of
 CC antibacterial drugs. The BASB034 products of the invention can be used
 CC screen for new antibacterial compounds that may target resistant
 CC bacteria.
 XX
 SQ Sequence 442 AA;

Query Match 100.0%; Score 2360; DB 21; Length 442;
 Best Local Similarity 100.0%; Pred. No. 2.8e-219;
 Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKVSLSTLTLSILPCFALLATQQAQVNPVAFVDEVRSKNDLQDNEILLIGVQSATQSA 60
 Db 1 mkvslstltlsilpcfallatqgaqvpnpvafvdevrskndlgdneilligvqsatqsa 60
 QY 61 STDTPANPLDEHEPELYTTALENKTMLNCSALNODIMRLACYDPLVHGETPAVTKTKRSI 120
 Db 61 stdtanpldehepelyttalenktmnlncsalnqdimrlacydplvhetpavtktkrsi 120
 QY 121 RLDETIWQTIKGKQVYQETTPDIFLMGNEKGMMLTKKDAKQLEYAAKQFTPLSFLDLD 180
 Db 121 rldetiwtikgkqvvyqettpdiflmgnkmgmltkkdaqlkleyaakqftplsflsfdld 180
 QY 181 RNNTPPLSSRRPHNPMYVLPIMHGKPNRSPNTPSHEARQFTPNFRAPELKFQVSVKVA 240
 Db 181 rnnptplsrrphnmpmyvlpimhkgpnrsnptpshearqftpnfrapelekfvsvkvka 240
 QY 241 AEDLWGTSDSLWFGVYQSQSHQIFNGKNSRPRFVRHDYQPEIFLQPVYSDLPWQKVRMI 300
 Db 241 aedlwgtstdslwfgvtqsqshqifngknsrprfrvhdypelilcpvyysdlpwqgkvrmi 300
 QY 301 GMGAVHSHNGESAKLSRSNRYAYLMAGMEWKNLTVMPRINGRIEKGSGSQPDNDPDLID 360
 Db 301 gmgavhshngesaklsrswnraylmagmewknlvtmpringrifksgsgsqpdndpdlid 360
 QY 361 YYGVDVRFYQLENKSNISTVRYNPRSGKALQLODYVPLGKIGSYFOIFOGYQGS 420
 Db 361 yygvgdvrflyqlenksnigtvrynprsgkqalqldyvplygkigsgyfgifggygqsl 420
 QY 421 IDYNHEATSFVGWGLNDWMGL 442
 Db 421 idynheatsfvgwglndwmgl 442

RESULT 2
 AAY85271
 ID AAY85271 standard; Protein; 442 AA.
 XX
 AC AAY85271;
 XX

DT 29-JUN-2000 (first entry)
 XX BASB034 amino acid sequence #4.
 DE
 XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;
 KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
 KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
 KW hearing loss; antibacterial drug.
 XX
 OS Moraxella catarrhalis.
 XX WO200015802-A1.
 PN
 XX 23-MAR-2000.
 PD
 XX 14-SEP-1999; 99WO-EP06781.
 PF
 XX 14-SEP-1998; 98GB-0020002.
 PR
 XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 PA
 XX Ruelle J;
 PI
 XX WPI; 2000-271440/23.
 DR N-PSDB; AAA10703.
 DR
 XX Novel BASB034 polynucleotides and polypeptides from Moraxella
 PT catarrhalis used to prepare vaccines against bacterial infections
 PT
 XX Claim 3; Page 69; 106pp; English.
 PS
 XX This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 CC strain Mc2969. The invention relates to BASB034 polypeptides from
 CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
 CC polynucleotides and polypeptides may be employed as research reagents and
 CC material for the discovery of treatments and diagnostics for diseases,
 CC particularly human diseases. They are particularly used to diagnose and
 CC treat M. catarrhalis infections. They can be used for diagnosis of
 CC disease, staging of disease, or determining response of an infectious
 CC organism to drugs. The polynucleotides may be used as a source for
 CC hybridization probes, and for screening of genetic mutations, serotype,
 CC organism or strain identification, identification of mutations in BASB034
 CC sequences, and as components of arrays which are useful for diagnostic
 CC and prognostic purposes. The polypeptides can be used to produce
 CC antibodies. The polypeptides can also be used in vaccine formulations,
 CC and to identify agonists and antagonists. The polypeptides, antibodies,
 CC agonists and antagonists (which are bacteriostatic) are used for the
 CC treatment and prevention of diseases such as otitis media in infants and
 CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
 CC invasive diseases, and chronic otitis media with hearing loss. The
 CC polypeptides, agonists and antagonists are also used for screening of
 CC antibacterial drugs. The BASB034 products of the invention can be used
 CC screen for new antibacterial compounds that may target resistant
 CC bacteria.
 XX
 SQ Sequence 442 AA;

Query Match 99.18; Score 2339; DB 21; Length 442;
 Best Local Similarity 99.18; Pred. No. 3e-217;
 Matches 438; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MKVSLSTLTLSILPCFALLATQQAQVNPVAFVDEVRSKNDLQDNEILLIGVQSATQSA 60
 Db 1 mkvslstltlsilpcfallatqgaqvpnpvafvdevrskndlgdneilligvqsatqsa 60
 QY 61 STDTPANPLDEHEPELYTTALENKTMLNCSALNODIMRLACYDPLVHGETPAVTKTKRSI 120
 Db 61 stdtanpldehepelyttalenktmnlncsalnqdimrlacydplvhetpavtktkrsi 120
 QY 121 RLDETIWQTIKGKQVYQETTPDIFLMGNEKGMMLTKKDAKQLEYAAKQFTPLSFLDLD 180
 Db 121 rldetiwtikgkqvvyqettpdiflmgnkmgmltkkdaqlkleyaakqftplsflsfdld 180

QY 181 RNTPLWSSRPHNPMYVLPFIHMGKPNRSPNTPSHEARQFTPNEFRAPELKFQVSVKVA 240
|||||
Db 181 rntplwssrphnmpvlpfihmgkpnrsntpsheakftpnfrapelkfqsvkvka 240
QY 241 AEDLWGTSDSLWFGYTOQSHWQIFNGKNRPRVRHDYQPEIFLTQPVYSDLPWDGKVRMI 300
|||||
Db 241 aedlwgtstdslwfgytqgshwqifngknsrprvrhdypqeilftqpvysdlpwwdgkvrmi 300
QY 301 GMGAVHSHNGESAKLSRSWNRAYLMAGMEWKNLTVMPTWGRIFKFGSGSQPDNDPDIID 360
|||||
Db 301 gmgavhshngesaksrswnraylmagmewknlvtmptwgrifkfgsgsqdpndpdiid 360
QY 361 YYGVDVRFYOLENKSNIISGTVRYNPRSGKALQDLYVPLGKISGYSYFQIFQGYGQSL 420
|||||
Db 361 yygvgdvrflyqlenksnlsigtvrynpnrgkqalqldyvplgkigsgyfifqgygqsl 420
QY 421 IDYNHEATSFYGLMLNDWMGL 442
|||||
Db 421 idynheatsfgyglmlndwmgl 442
RESULT 3
AA85268
ID AAY85268 standard; Protein; 442 AA.
XX AC
XX AAY85268;
DT 29-JUN-2000 (first entry)
XX DE
XX BAS034 amino acid sequence #1.
KW Moraxella catarrhalis infection; BAS034; diagnosis; staging;
KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW hearing loss; antibacterial drug.
XX OS
XX Moraxella catarrhalis.
XX PN
XX WO200015802-A1.
PD 23-MAR-2000.
XX PF
XX 14-SEP-1999; 99WO-EP06781.
XX PR
XX 14-SEP-1998; 98GB-0020002.
XX PA
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
PI Ruelle J;
XX DR
XX WPI; 2000-271440/23.
XX DR
XX N-PSDB; AAA10700.
PT Novel BAS034 polynucleotides and polypeptides from Moraxella
XX catarrhalis used to prepare vaccines against bacterial infections
PS Claim 3; Fig 2; 106pp; English.
XX
This sequence represents a Moraxella catarrhalis BAS034 polypeptide from
strain Mc2931 (ATCC 43617). The invention relates to BAS034 polypeptides
from M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The
BAS034 polynucleotides and polypeptides may be employed as research
reagents and material for the discovery of treatments and diagnostics for
diseases, particularly human diseases. They are particularly used to
diagnose and treat M. catarrhalis infections. They can be used for
diagnosis of disease, staging of disease, or determining response of an
infectious organism to drugs. The polynucleotides may be used as a source
for hybridization probes, and for screening of genetic mutations,
serotype, organism or strain identification, identification of mutations
in BAS034 sequences, and as components of arrays which are useful for
diagnostic and prognostic purposes. The polypeptides can be used to
produce antibodies. The polypeptides can also be used in vaccine

CC formulations, and to identify agonists and antagonists. The polypeptides,
CC antibodies, agonists and antagonists (which are bacteriostatic) are used
CC for the treatment and prevention of diseases such as otitis media in
CC infants and children, pneumonia in elderlies, sinusitis, nosocomial
CC infections and invasive diseases, and chronic otitis media with hearing
CC loss. The polypeptides, agonists and antagonists are also used for
CC screening of antibacterial drugs. The BAS034 products of the invention
CC can be used screen for new antibacterial compounds that may target
CC resistant bacteria.
XX
SQ Sequence 442 AA;
Query Match 98.7%; Score 2330; DB 21; Length 442;
Best Local Similarity 98.6%; Pred. No. 2.2e-216;
Matches 436; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 MKVSLTLTSLILPCFALIAIOQAQAVPNPVAFAVDEVRSKNDLQDNELLIGVOSTQSA 60
|||||
Db 1 mkvslstltlslilscfaillaigqagavpnpvafdevrsendlgqdnelpidvqsatqsa 60
QY 61 STDTANPLDEHEPELYTTALENKTMLINCSALNQDITMLRACYDTLVHGETPAVIKTKRSI 120
|||||
Db 61 stdtanpldehepelyttalenktmlinclsalnqdmrlacydtlvhgetpaviktkrsi 120
QY 121 RLDETIWQTIKGPVVOYQETTPFIPLMGNEKGMFTKDAKOLEYAAKQFTPLSLSFOLD 180
|||||
Db 121 rldetiwtikgkpvvyoyettppifplmgnekgmftkdaqkqleyaakqftplslsfold 180
QY 181 RNTPLWSSRPHNPMYVLPFIHMGKPNRSPNTPSHEARQFTPNEFRAPELKFQVSVKVA 240
|||||
Db 181 rntplwssrphnmpvlpfihmgkpnrsntpsheakftpnfrapelkfqsvkvka 240
QY 241 AEDLWGTSDSLWFGYTOQSHWQIFNGKNRPRVRHDYQPEIFLTQPVYSDLPWDGKVRMI 300
|||||
Db 241 aedlwgtstdslwfgytqgshwqifngknsrprvrhdypqeilftqpvysdlpwwdgkvrmi 300
QY 301 GMGAVHSHNGESAKLSRSWNRAYLMAGMEWKNLTVMPTWGRIFKFGSGSQPDNDPDIID 360
|||||
Db 301 gmgavhshngesaksrswnraylmagmewknlvtmptwgrifkfgsgsqdpndpdiid 360
QY 361 YYGVDVRFYOLENKSNIISGTVRYNPRSGKALQDLYVPLGKISGYSYFQIFQGYGQSL 420
|||||
Db 361 yygvgdvrflyqlenksnlsigtvrynpnrgkqalqldyvplgkigsgyfifqgygqsl 420
QY 421 IDYNHEATSFYGLMLNDWMGL 442
|||||
Db 421 idynheatsfgyglmlndwmgl 442
RESULT 4
AA85270
ID AAY85270 standard; Protein; 442 AA.
XX AC
XX AAY85270;
DT 29-JUN-2000 (first entry)
XX DE
XX BAS034 amino acid sequence #3.
KW Moraxella catarrhalis infection; BAS034; diagnosis; staging;
KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW hearing loss; antibacterial drug.
XX OS
XX Moraxella catarrhalis.
XX PN
XX WO200015802-A1.
PD 23-MAR-2000.
XX PF
XX 14-SEP-1999; 99WO-EP06781.
XX

PR 14-SEP-1998; 98GB-00200002.
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX Ruelle J;
XX WPI; 2000-271440/23.
DR N-PSDB; AAA10702.
XX
PT Novel BASB034 polynucleotides and polypeptides from Moraxella
PT catarrhalis used to prepare vaccines against bacterial infections
XX
PS Claim 3; Page 68; 106pp; English.
XX
XX This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC strain Mc2913. The invention relates to BASB034 polypeptides from
CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
CC polynucleotides and polypeptides may be employed as research reagents and
CC material for the discovery of treatments and diagnostics for diseases,
CC particularly human diseases. They are particularly used to diagnose and
CC treat M. catarrhalis infections. They can be used for diagnosis of
CC disease, staging of disease, or determining response of an infectious
CC organism to drugs. The polynucleotides may be used as a source for
CC hybridization probes, and for screening of genetic mutations, serotype,
CC organism or strain identification, identification of mutations in BASB034
CC sequences, and as components of arrays which are useful for diagnostic
CC and prognostic purposes. The polypeptides can be used to produce
CC antibodies. The polypeptides can also be used in vaccine formulations,
CC and to identify agonists and antagonists. The polypeptides, antibodies,
CC agonists and antagonists (which are bacteriostatic) are used for the
CC treatment and prevention of diseases such as otitis media in infants and
CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
CC invasive diseases, and chronic otitis media with hearing loss. The
CC polypeptides, agonists and antagonists are also used for screening of
CC antibacterial drugs. The BASB034 products of the invention can be used
CC screen for new antibacterial compounds that may target resistant
CC bacteria.
XX
SQ Sequence 442 AA;

Query Match 98.7%; Score 2330; DB 21; Length 442;
Best Local Similarity 98.9%; Pred. No. 2.2e-216;
Matches 437; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKVSLSTLTLSILPCFATLATOQAQVNPVAFVDEVRSKNDLGDNEELLGVQSAQSA 60
DB 1 mkvslstltlsilscfallalqqakavnpvafvdevrsendlqgdneipldvgsatqsa 60
QY 61 STDTPANPLDEHEPELYTTALENKTMLNCSALNQDIMRLACYDTLVHGETPAVTKRSI 120
DB 61 stdtanpldehepelyttalenktmlncsalnqdimrlacydtlvhgetpavtkrksi 120
QY 121 RLDETWTQITKGQVYQETDIFLGMNKGKMLTKDAKQLEYAAKQFTPLSLSPDL 180
DB 121 rldetwtqitkgqvvyqetdipflgmnekgmltkdakqleyaakqftplslsfdld 180
QY 181 RNNTPLWSSRPHNPMYVLPIMHCKPNRSPNTPSHEARQFTPNFRAPELKFQVSVKVA 240
DB 181 rnntplwssrphnpyvlpimfhgkpnrsntpshearqftpnfrapelkfqvsvkva 240
QY 241 AEDLWGTDSLWFGYQQSHWQIFNGKNSRPFVRVHDYQPEIFLTQPVYSDLPDCKVRMI 300
DB 241 aedlwgtdslwfgytqshwqifngknsrpfvrhdyqpeifltqpvysdldpdckvrmi 300
QY 301 GMGAVHNSGESAKLSRSWNRAYLMAGNEWKNLTVMPRIWGRIFKESGSGQPDNDPILD 360
DB 301 gmgavhnsngesaklsrswnraylmagnewknltvmpriwgrifkessgsgqpdndpild 360
QY 361 YYGVDVRFVLYOLENKNISGTVRNPNSRGKALQLDVYVPLGKGISGYFOFGYQSL 420
DB 361 yygvgdvrfvlyolenknisgtrvnpnsrgkalqldvyvplgkgisgyfifggyqgsi 420

QY 421 IDYNHEATSGVGLMLNDWMGL 442
DB 421 idynheatsfgvglmndwmgl 442
RESULT 5
AA75156
ID AAY75156 standard; Protein; 370 AA.
XX
AC AAY75156;
XX 21-MAR-2000 (first entry)
XX
XX Neisseria gonorrhoea ORF 582 protein sequence SEQ ID NO:1786.
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy.
XX
OS Neisseria gonorrhoeae.
XX
PN WO9957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.
XX
PR 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
XX (CHIR) CHIRON CORP.
XX (GENO-) INST GENOMIC RES.
XX
XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
DR WPI; 2000-062150/05.
DR N-PSDB; AA253918.
XX
XX Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics -
XX
PS Claim 2; Page 903; 1453pp; English.
XX
XX AA253015 to AA254536, AA254577 to AA254615, and AA74253 to AA75941
CC represent novel Neisseria meningitis and N. gonorrhoea polynucleotides
CC and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
SQ Sequence 370 AA;

Query Match 34.8%; Score 822; DB 21; Length 370;
Best Local Similarity 44.7%; Pred. No. 8.5e-71;
Matches 163; Conservative 68; Mismatches 110; Indels 24; Gaps 8;

QY 87 INCSALNODIMRLACYDTLVHGETPAVI-----KTKRSIRLDETWTQIT-KGKPVVYQOE 140


```

Query Match      10.4%  Score 246.5;  DB 19;  Length 355;
Best Local Similarity 25.1%  Pred. No. 3.2e-15;
Matches 82;  Conservative 41;  Mismatches 117;  Indels 87;  Gaps 11

```

QY 217 ARQFTNEFRAPELKQVSVKVAEDLWGTDSDLWFGYTQOSHWOIFNGKNSRPFVRVD 276
Db 103 ---yqrnef-----kqisfrvprfhrilwtktlylaytqtdwfdyndpqapmnmn 154
QY 277 YQPEIFLTQPVYSDLPWDGKV---RMIGCAVHHSNG-ESAKLSRSWNRAYLMAGMEWKN 332
Db 155 fmpellyvypf-nfkpgfgkignfseiwqhinsvggagcyqpfnk-----egnpeng 209
QY 333 LTVMPRI-----WGRIKFGSGSQP-----DDNDP 357
Db 210 fpgqvivdyngqgkdvrvwgcsvsagrpvfrlvwekggklmwaywpyvpydgsnnpn 269
QY 358 ILDYGYGVDVRFLY-----QLENKSNISGTVRYNPRSGKGALQLDYVYPLGKISGYF 410
Db 270 lidymygnakidyrgrhhfelqllydiftqwyrd--rwhgafrigytyrinpfvgiya 327
QY 411 QIFQGYGQSLIDYNHEATSGVGLMLN 437
Db 328 qwingydglyeydvfsnrigvgirln 354

RESULT 11

AA10960
ID AAY10960 standard; Protein; 356 AA.

XX
AC AAY10960;

DT 08-JUN-1999 (first entry)

DE H. pylori ORF 07ap80601_5083193_f3_8 cell envelope protein.

XX Vaccine; probe; diagnostic; ORF; cell envelope protein;

KW secreted protein; cellular protein.

XX Helicobacter pylori.

XX WO9818323-A1.

PD 07-MAY-1998.

PF 28-OCT-1997; 97WO-US19575.

PR 14-JUL-1997; 97US-0891928.

PR 28-OCT-1996; 96US-0739150.

PR 06-DEC-1996; 96US-0759739.

XX (ASTR) ASTRA AB.

XX Alm RA, Smith D;

XX WPI; 1998-271811/24.

XX N-PSDB; AAX30427.

XX Helicobacter pylori nucleic acids and proteins - used to develop
PT products for the detection, prevention and treatment of H. pylori
PT infections

XX Claims 27, 31; Page 160-161; 279pp; English.

XX Recombinant or substantially pure preparations of H. pylori polypeptides
CC are disclosed, together with the nucleic acids encoding them. In all,
CC 73 ORFs are shown. The proteins are variously cell envelope proteins,
CC secreted proteins or other cellular proteins. Vaccines containing the
CC nucleic acids or proteins are claimed, as are probes containing at least
CC 8 nucleotides from the nucleic acid sequences. The vaccines are useful
CC for treating or reducing the risk of H. pylori infections, and the
CC probes can be used diagnostically for detecting the presence of
CC Helicobacter in a sample. The products are also of use in screening
CC for compounds having the ability to interfere with the H. pylori life
CC cycle or to inhibit H. pylori infection.

XX Sequence 356 AA;

SQ

Query Match 10.1%; Score 239.5; DB 19; Length 356;
Best Local Similarity 24.8%; Pred. No. 1.5e-14;
Matches 82; Conservative 43; Mismatches 111; Indels 95; Gaps 12;
QY 157 KKDAKOLEYAAKQFTPLSLSFDDLDRNNTPLWSSRPHNPMYVLFIEMHGKPNRSPNTPSHE 216
Db 70 kkylnmmdylgtlylfpfyhaf-----tpifqwyhpnlnp----- 103
QY 217 ARQFTNEFRAPELKQVSVKVAEDLWGTDSDLWFGYTQOSHWOIFNGKNSRPFVRVD 276
Db 104 ---yqrnef-----kqisfrvprfhrilwtktlylaytqtdwfdyndpqapmnmn 155
QY 277 YQPEIFLTQPVYSDLPWDGKV---RMIGCAVHHSNG-ESAKLSRSWNRAYLMAGMEWKN 332
Db 155 fmpellyvypf-nfkpgfgkignfseiwqhinsvggagcyqpfnkengnpengpfgq 214
QY 323 -----YLMAGMEWK---NLTVMNPRIMGRIFKEGSGSQPDD- 354
Db 215 pviavdyngqgkdvrvwgcsvsxxgnxlcflvwekggklmwaywpyv-----pydq 266
QY 355 -NPDILDYGYGVDVRFLY-----QLENKSNISGTVRYNPRSGKGALQLDYVYPLGKGI 406
Db 267 snpqlidymygnakidyrgrhhfelqllydiftqwyrd--rwhgafrigytyrinpfv 324
QY 407 SGYFQIFQGYGQSLIDYNHEATSGVGLMLN 437
Db 325 giyaqwingydglyeydvfsnrigvgirln 355

RESULT 12

AAW20760

ID AAW20760 standard; protein; 253 AA.

XX
AC AAW20760;

DT 15-JUL-1997 (first entry)

DE H. pylori outer membrane protein, 07ap80601orf8.

XX Cytoplasmic; vaccine; prevention; infection; treatment; identification;
KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope;
XX outer membrane.

OS Helicobacter pylori.

XX WO9640893-A1.

XX 19-DEC-1996.

XX 06-JUN-1996; 96WO-US09122.

XX 01-APR-1996; 96US-0630405.

XX 07-JUN-1995; 95US-0487032.

XX (ASTR) ASTRA AB.

XX Berglindh OT, Smith D, Mellgaerd BL;

XX WPI; 1997-052306/05.

XX N-PSDB; AAT68013.

XX Helicobacter pylori nucleic acid sequences and related
PT polypeptides(s) - useful for vaccines to treat or prevent H. pylori
PT infection, and to detect Helicobacter

XX Claim 56; Page 1172-1173; 1481pp; English.

XX The present sequence is a Helicobacter pylori outer membrane protein.

CC The protein may be used in a vaccine to prevent or treat H. pylori

CC infection or to identify H. pylori polypeptide binding compounds, useful

CC as potential H. pylori life cycle activators or inhibitors. The genomic


```

QY      47 NELLIGVQSASTSDTANPLDEHEPELYTTALENKTMILNCALNQDIMRLACVDFTLV 100
|       |   ||         |   |   |   |   |   |   |   |   |   |   |
Db     525 NHCLGSGSHEHIONCKKWDI-----ETKLQLLTLSAMCQNLAETAEDD--- 565

QY      107 HGTPAVIKTKRSIRLDETITWOTIKGPOVVYQOETTDPF----- 146
|       |   |: :||    |: :||    |: :||    |: :||    |: :||    |: :||
Db     566 --EAP-----VDLNKYLIQEKPKEVMTRNHPVEELDHYHQVELAQENQHRAV 615

QY      147 -----LMNGEKMLTKDAKALEYAA-----KQTPTLSFDLDNRNTPLWS 188
|       |   |: :||    |: :||    |: :||    |: :||    |: :||    |: :||
Db     616 DQVIKAVRKICSALDGVEIPSVTEA-VKKLKRAVLNPRKNSADVTSLSGSDTRKNSTKGS 674

QY      189 SRPHNPMYULP-----IPMHOKPNR----SPNTPSHBEAROPTNFEPRAPELKFVOV 234
|       |   |: :||    |: :||    |: :||    |: :||    |: :||    |: :||
Db     675 LNPNPVQVSMDHLTRIYDLLRLHANSSRCSTGCCPRGRNIKEAWTATE-----OL 726

QY      235 SVKVKAEDL--WGTDSDLWFYGTYTOOSHWQIFNGKN-SRPFP---RVHDYQPEIFLTQ-- 285
|       |   |: :||    |: :||    |: :||    |: :||    |: :||    |: :||
Db     727 QFTVYAAGHGSISSNWVNSEKYIYLICLSH-----NGKDLFKPIQSKVKGYTNFFYLIKWD 782

QY      286 ----PV-YSDLPWDGVKMGMGAHVHHSNGESAKLRSWN-----RAYL 324
|       |   |: :||    |: :||    |: :||    |: :||    |: :||    |: :||
Db     783 ELIIPIQISQPLESVLHTLFGLNLQGSSGPSDKORKGPALKVYSLTLEDFKRFL 842

QY      325 MAGMEKNLTVMPRINGRIKFGSGSQPDNDPIDLYGGDVRFYLYQLENKSINSGTGTVR 384
|       |   |: :||    |: :||    |: :||    |: :||    |: :||    |: :||
Db     843 TCG-----TKLLYLW-----TSSTNSIPGAIPKKSIVMERIVLQVDFPSPAFDIITY 889

QY      385 YNPRSCKGALOQDYVPYPLKGISGYFIQQGYCOSLIDYNHEATSRGV 432
|       |   |: :||    |: :||    |: :||    |: :||    |: :||    |: :||
Db     890 TSPOIDRNIIQQDKLETLESIDKG-----KLLDIHRDSFGL 927


RESULT          2
US-09-170-996-30
; Sequence 30, Application US/09170996
; Patent No. 6291220
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Molz, Lisa
; APPLICANT: Chen, Yen-Wen
; TITLE OF INVENTION: No. 6291220el PI 3-Kinase Polypeptides
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESSES:
; ADDRESSER: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/170,996
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/609,049
; FILING DATE: 29-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-063700US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1726 amino acids
; TYPE: amino acid

```

NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-0637000S
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1658 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-609-049A-13

Query Match 4.2%; Score 99.5; DB 2; Length 1658;
Best Local Similarity 18.6%; Pred. No. 0.65;
Matches 98; Conservative 79; Mismatches 190; Indels 161; Gaps 21;

Qy 1 MKVSLS-----TLTSLPCFAILAIQQAQAVNPVAFVD-----EVRKNDLGQD 46
Db 397 KVSIEIEGLQLPVTFCDVSVTEIIMQALCWHDLDNQVDVGSYILKVGQGEVLQN 456
Qy 47 NELLIGVSATQSASTDTANPLDEHEPELYTTALENKTMLNCSALNODIMRLACYDTLV 106
Db 457 NHCIGSHEHIQNCRKWD-----EIKQLLTLSAMCONLARTAEDD--- 497
Qy 107 HGETPAVIKTKRSIRLDETIWQTIKGKPVVYQETDPIF----- 146
Db 498 --EAP-----VLDNKLYQIEKPYKEVMIRHPVEELDSYHYQVELALQTNQHRV 547
Qy 147 -----LMGNEKGLTKKDAKOLEYAA-----KQFTPLSLSFDDLRNNTPLWS 188
Db 548 DQVIKAVRKICSGALDGVETPSVTEA-VKKLRVNLPRNKSADVTSLGSDTRKNSKGS 606
Qy 189 SRPNPMYVLP-----IFMHGKPNR-----SPNTPSHEARQFTNPEPAPELAFQV 234
Db 607 LNPNPQVQSMDDLHTTAYDILLRLHANSRSCSTGCPGRSRNIKEAWTATE-----QL 658
Qy 235 SVKVKAAEDL---WGTSDSLWFGYTTQSHWQIFNGKN-SRPF---RVHDOPEIFLQ-- 285
Db 659 QFTVYAAHGISSNWNVSNEYKYILCSLSH-----NGKDLFKPIQSKKVGTYKNFFYLKWD 714
Qy 286 ----PV-YSDLPWDGKVRMIGMAVHNSGESAKLSRSWN-----RAYL 324
Db 715 ELIIFPIQISOLPLESVLHLTLFGLVNLQSSGSPDSNQRKGPEALGVSLTLFDFKRF 774
Qy 325 MAGMEWKNLTVMPRIWGRIFKEGSGSQPDNDPILDYGYGDVRFYOLENKNISGTVR 384
Db 775 TCG-----TKLLYLW-----TSSHTNSIPGAIPKKSVMERIVLQVDFPSPAFDIY 821
Qy 385 YNPRSGKALQDLYVYPLGKIGSYFQIFQGYGQSLIDYNHEATSFV 432
Db 822 TSPQIDRNIIOQDKLETLESIDKG-----KLLDIIHRDSSFGL 859

RESULT 4
US-09-170-996-13
Sequence 13, Application US/09170996
Patent No. 6291220
GENERAL INFORMATION:
APPLICANT: Williams, Lewis T.
APPLICANT: Molz, Lisa
APPLICANT: Chen, Yen-Wen
TITLE OF INVENTION: No. 6291220el PI 3-Kinase Polypeptides
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA

ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/170.996
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/609,049
FILING DATE: 29-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-0637000S
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1658 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-170-996-13

Query Match 4.2%; Score 99.5; DB 4; Length 1658;
Best Local Similarity 18.6%; Pred. No. 0.65;
Matches 98; Conservative 79; Mismatches 190; Indels 161; Gaps 21;

Qy 1 MKVSLS-----TLTSLPCFAILAIQQAQAVNPVAFVD-----EVRKNDLGQD 46
Db 397 KVSIEIEGLQLPVTFCDVSVTEIIMQALCWHDLDNQVDVGSYILKVGQGEVLQN 456
Qy 47 NELLIGVSATQSASTDTANPLDEHEPELYTTALENKTMLNCSALNODIMRLACYDTLV 106
Db 457 NHCIGSHEHIQNCRKWD-----EIKQLLTLSAMCONLARTAEDD--- 497
Qy 107 HGETPAVIKTKRSIRLDETIWQTIKGKPVVYQETDPIF----- 146
Db 498 --EAP-----VLDNKLYQIEKPYKEVMIRHPVEELDSYHYQVELALQTNQHRV 547
Qy 147 -----LMGNEKGLTKKDAKOLEYAA-----KQFTPLSLSFDDLRNNTPLWS 188
Db 548 DQVIKAVRKICSGALDGVETPSVTEA-VKKLRVNLPRNKSADVTSLGSDTRKNSKGS 606
Qy 189 SRPNPMYVLP-----IFMHGKPNR-----SPNTPSHEARQFTNPEPAPELAFQV 234
Db 607 LNPNPQVQSMDDLHTTAYDILLRLHANSRSCSTGCPGRSRNIKEAWTATE-----QL 658
Qy 235 SVKVKAAEDL---WGTSDSLWFGYTTQSHWQIFNGKN-SRPF---RVHDOPEIFLQ-- 285
Db 659 QFTVYAAHGISSNWNVSNEYKYILCSLSH-----NGKDLFKPIQSKKVGTYKNFFYLKWD 714
Qy 286 ----PV-YSDLPWDGKVRMIGMAVHNSGESAKLSRSWN-----RAYL 324
Db 715 ELIIFPIQISOLPLESVLHLTLFGLVNLQSSGSPDSNQRKGPEALGVSLTLFDFKRF 774
Qy 325 MAGMEWKNLTVMPRIWGRIFKEGSGSQPDNDPILDYGYGDVRFYOLENKNISGTVR 384
Db 775 TCG-----TKLLYLW-----TSSHTNSIPGAIPKKSVMERIVLQVDFPSPAFDIY 821
Qy 385 YNPRSGKALQDLYVYPLGKIGSYFQIFQGYGQSLIDYNHEATSFV 432
Db 822 TSPQIDRNIIOQDKLETLESIDKG-----KLLDIIHRDSSFGL 859

RESULT 5
US-07-717-331F-2


```

;
; TYPE: amino acid
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: rat protein
; US-08-483-322-5

Query Match      3.7%; Score 87; DB 1; Length 355;
Best Local Similarity 19.3%; Pred. No. 1.1;
Matches 59; Conservative 36; Mismatches 97; Indels 114; Gaps 16;

QY 45 QNELLIG--VOSATQSASTDTA-NPLDHEPELYTALLENKTMNLINCSALNODIMRLAC 101
   :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:
Db 7  RDAAFFIGTSLATSTESNTNPTGKPKNEENEDE-----TDKYPNFGSGIDDD-----ED 56

QY 102 YDTLVHGETPAV-IKTKRSIRLDETIWQIKGPKQVYVQETDPIFLMGNEKGLTKKDA 160
   :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:
Db 57 FSISSITATTPWSAHTKQN--QERTQWNPISNPEVLLQTTT-----96

QY 161 KOLEYAAKQFTPLSLSFOLDNRNNTPL-----WSSRPHNPMYVLPFIEMHGKPNRSPNTPSHE 216
   :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:
Db 97 -----RMT-----DIDRNSTSAHGENWTQEPQPPF-----NNHEVQ 127

QY 217 ARQFTNEFRAPELKFOVSVKVAEDLM-----GTSDLMFGYGTQSHWQIFNG 266
   :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:
Db 128 DEEETPH-----ATSTTWADPNSTTEEAATQKEKWF-----ENEMQ--G 164

QY 267 KN-----SRPFRVHDYQPEIFLTQPVYSDLPWDGKVRMIG--MGAVHHNGES 312
   :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:
Db 165 KNPPTPSEDSHVTEGTASAHNNHPSQRMSTQSQEDVSWTDFDPIISHPMGQGHQTESKD 224

QY 313 AKLSRS 318
   |
Db 225 TGSSES 230

RESULT 9
US-08-483-322-5
; Sequence 5, Application US/08483322
; Patent No. 5760178
; GENERAL INFORMATION:
; APPLICANT: HERRLICH, Peter
; APPLICANT: PONTA, Helmut
; APPLICANT: GUENTHER, Ursula
; APPLICANT: MATZKU, Siegfried
; APPLICANT: WENZL, Achim
; TITLE OF INVENTION: VARIANT CD44 SURFACE PROTEINS, DNA
; TITLE OF INVENTION: SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
; TITLE OF INVENTION: AS WELL AS THEIR USE IN DIAGNOSIS AND THERAPY
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,322
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/946,497
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16915/145
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids

```

```

;
; TYPE: amino acid
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: rat protein
; US-08-483-322-5

Query Match      3.7%; Score 87; DB 1; Length 355;
Best Local Similarity 19.3%; Pred. No. 1.1;
Matches 59; Conservative 36; Mismatches 97; Indels 114; Gaps 16;

QY 45 QNELLIG--VOSATQSASTDTA-NPLDHEPELYTALLENKTMNLINCSALNODIMRLAC 101
   :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:
Db 7  RDAAFFIGTSLATSTESNTNPTGKPKNEENEDE-----TDKYPNFGSGIDDD-----ED 56

QY 102 YDTLVHGETPAV-IKTKRSIRLDETIWQIKGPKQVYVQETDPIFLMGNEKGLTKKDA 160
   :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:
Db 57 FSISSITATTPWSAHTKQN--QERTQWNPISNPEVLLQTTT-----96

QY 161 KOLEYAAKQFTPLSLSFOLDNRNNTPL-----WSSRPHNPMYVLPFIEMHGKPNRSPNTPSHE 216
   :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:
Db 97 -----RMT-----DIDRNSTSAHGENWTQEPQPPF-----NNHEVQ 127

QY 217 ARQFTNEFRAPELKFOVSVKVAEDLM-----GTSDLMFGYGTQSHWQIFNG 266
   :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:
Db 128 DEEETPH-----ATSTTWADPNSTTEEAATQKEKWF-----ENEMQ--G 164

QY 267 KN-----SRPFRVHDYQPEIFLTQPVYSDLPWDGKVRMIG--MGAVHHNGES 312
   :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:
Db 165 KNPPTPSEDSHVTEGTASAHNNHPSQRMSTQSQEDVSWTDFDPIISHPMGQGHQTESKD 224

QY 313 AKLSRS 318
   |
Db 225 TGSSES 230

RESULT 10
US-08-478-882-5
; Sequence 5, Application US/08478882
; Patent No. 585575
; GENERAL INFORMATION:
; APPLICANT: HERRLICH, Peter
; APPLICANT: PONTA, Helmut
; APPLICANT: GUENTHER, Ursula
; APPLICANT: MATZKU, Siegfried
; APPLICANT: WENZL, Achim
; TITLE OF INVENTION: VARIANT CD44 SURFACE PROTEINS, DNA
; TITLE OF INVENTION: SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
; TITLE OF INVENTION: AS WELL AS THEIR USE IN DIAGNOSIS AND THERAPY
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,882
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/946,497
; FILING DATE: 19921109
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16915/145
; TELECOMMUNICATION INFORMATION:

```


APPLICATION NUMBER: 9321301.5
FILING DATE: 15-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: DY007.001APC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714-760-0404
TELEFAX: 714-760-9502
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1088 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-633-768A-1

Query Match 3.6%; Score 86; DB 3; Length 1088;
Best Local Similarity 21.6%; Pred. No. 8.5;
Matches 72; Conservative 40; Mismatches 138; Indels 84; Gaps 15;

QY 106 VHGEPAY-IKTKRSIRIDETIMQIKKQVYVQETDPIFLMGNEKMLTKKAKOLE 164
DB 29 IHSAPPAVATARKTNRLN-VSMTALSDKQATAGSTD-----NPDGI---DKTYD 76
QY 165 YAAK-QETPLS-----LSFDLRNNTPLWSSRRHNPVYLPIM 202
DB 77 YGVGAGFSLSTNMFAGSSITPGITDWTATMNNFDRIDNP--SITVQHVVQVQVTSY 134
QY 203 HGKPRSPPTSGHEARQPTNEFFRAPELKVYVKAEDLMGDSIDMFGYTOOSHWO 262
DB 135 NNNSYRVFNPDPGPIRDYV---RGPILKQQLD-WIRQELSECCDPCG----- 178
QY 263 IFNGKNSRPERHDYQPELFLQPYVSDLPMDGKVRMIGMGAHVHNSGSAKLSNMRA 322
DB 179 TPTSGFLTFETKDLSTVITYGKFKTRVTRKSDGKVI-----ENDVGTAS----- 224
QY 333 YLMAGEMKNTLVMPRIMGRIFKEGSGSQDPDNPILDYVGYGVDFELTQ--LENKSN 379
DB 225 ---SGNKRGMLFVDRILGNALIASVKNKFNRAVAKQEGFYGAGVCKKQDIILERTOI 281
QY 380 SGT-----VRYN-----PRSGKALQLDYVPL 402
DB 282 AMTNVNYDNLNMQWDLRPPHHDGALNDPYIIPM 315

RESULT 15
US-09-091-117-5
Sequence 5, Application US/09091117
Patent No. 6171589
GENERAL INFORMATION:
APPLICANT: The University of Melbourne
TITLE OF INVENTION: Mycoplasma Recombinant Polypeptides and
TITLE OF INVENTION: Vaccines
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESSES:
ADDRESSEE: GREENLEE, WINNER and SULLIVAN P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: United States of America
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/091.117
FILING DATE: 12 JUNE 1998

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU96/00803
FILING DATE: 13-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PN7127
FILING DATE: 13-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: WINNER, Ellen P.
TELECOMMUNICATION INFORMATION:
TELEPHONE: +1 303 499 8080
TELEFAX: +1 303 499 8080
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1024 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Mycoplasma genitalium
US-09-091-117-5

Query Match 3.6%; Score 84.5; DB 4; Length 1024;
Best Local Similarity 20.9%; Pred. No. 11;
Matches 89; Conservative 60; Mismatches 170; Indels 107; Gaps 23;

QY 1 MKVVS-LSITLISILPCFAILAIQQAQAVNPVAFVDEVRKNDLGODNELLIGVOSATOS 59
DB 1 MKLSTYITLISISGAFGTAL---ALPTVALL-----KHHQOONTE----- 40
QY 60 ASTDPANPDEHEPELYTALENKTKMLNCSALNDIMRLACYDTLY-HGETPAVYIKTR 118
DB 41 ---KOONPKIDIRFGL-----NNVOVPPTIPRHQTVVENVTKNAVALVDYKAPQKFEAK 91
QY 119 SIRLDETIMQIKKQVYVQETDPIFLMGNEKMLTKKAKOLEYAAKO-FTPLSLSF 177
DB 92 S-----ALNNKIQVFEFK-----FL--RTGVYINLMDKEMIDQTLFIPQSF 135
QY 178 DLDNRNTPLMSSRPHPNMYLP-IFMHGKPNRSP-NTPSHEARQPTNEFFRAPELKVY 235
DB 136 DLSANKLNTLTSNQEVSILDEIFTFNFSDKNQPLKLPDGSVVVAVANE-----SYVYSVK 191
QY 236 VVYKAEDLMGDSIDMFGYTOOSHWOIFNGKNSRFRHDYQPEL-FLTQPYVSDLPW 293
DB 192 ATLQKLKLVLYSRADHSVIGISYAIPTVSLNGKTQNDSEFNPKSNINFAKKNVYALNPF 251
QY 294 DCKVRMIGMAVHHS--NGESAK-----LSRSWNRAYLMAGEMKNTLVMPRIMGRIFK 345
DB 252 EAOQYLVGGKFLNCKVKNADVDKNDINNHIETQFNVAKITA-----TLGKAKR 300
QY 346 ---EGSGSQP-----DNDPDILDY--GYGV-----RFLYQL-- 373
DB 301 QFGEHKNQGPLSLKLKLVLSGLNNEFKQLFNYVRPGLDFFVSDLIQSSSGSNKKTYVOLLF 360
QY 374 ENKSN 379
DB 361 ENKTYI 366

Search completed: November 30, 2001, 14:17:45
Job time: 222 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:10:38 ; Search time 72.04 Seconds
(without alignments)
454.475 Million cell updates/sec

Title: US-09-787-083-2

Perfect score: 2360

Sequence: 1 MKVSLTLTSLTSLCEAFLA.....YNHEATSPFGVGLNMDMGL 442

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_1101.*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2360	100.0	442	21	BA5B034 amino acid
2	2354	99.7	442	21	BA5B034 amino acid
3	2352	99.7	442	21	BA5B034 amino acid
4	2330	98.7	442	21	BA5B034 amino acid
5	821	34.8	370	21	Neisseria gonorrhoe
6	814	34.5	370	21	Neisseria meningit
7	814	34.5	370	21	Neisseria meningit
8	814	34.5	374	21	Neisseria meningit
9	811	34.4	375	21	Neisseria meningit
10	246.5	10.4	355	19	H. pylori GHP0 172
11	239.5	10.1	356	19	H. pylori ORF 07ap

12	156.5	6.6	253	18	AAW20760
13	117	5.0	1686	8	AAW70991
14	110.5	4.7	800	19	AAW70420
15	110.5	4.7	822	13	AAW26021
16	109	4.6	824	21	AAW23180
17	107	4.5	157	18	AAW20538
18	103	4.4	798	21	AAW40925
19	103	4.4	798	21	AAW90225
20	103	4.4	798	22	AAW83944
21	102.5	4.3	537	18	AAW91000
22	100.5	4.3	1726	18	AAW38756
23	98.5	4.2	467	20	AAW13378
24	98.5	4.2	467	21	AAW94860
25	98.5	4.2	467	22	AAW88401
26	98.5	4.2	467	22	AAW80246
27	96	4.1	888	22	AAW70751
28	95.5	4.0	502	22	AAW90031
29	95.5	4.0	516	22	AAW78876
30	95.5	4.0	682	17	AAW04359
31	95	4.0	1227	22	AAW81501
32	94	4.0	372	16	AAW69607
33	94	4.0	522	22	AAW95616
34	93.5	4.0	467	19	AAW75057
35	93	3.9	761	20	AAW99084
36	93	3.9	1717	22	AAW20498
37	92	3.9	600	22	AAW75092
38	92	3.9	1627	16	AAW67538
39	91.5	3.9	857	13	AAW2814
40	91.5	3.9	1024	18	AAW19604
41	91	3.9	564	21	AAW10457
42	91	3.9	564	21	AAW81952
43	90	3.8	522	22	AAW95513
44	90	3.8	790	22	AAW93045
45	90	3.8	841	20	AAW24318

ALIGNMENTS

RESULT 1
ID AAY85268 standard; Protein; 442 AA.
XX
AC AAY85268;
XX
DT 29-JUN-2000 (first entry)
XX
DE BASB034 amino acid sequence #1.
XX
KW Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW hearing loss; antibacterial drug.
XX
OS Moraxella catarrhalis.
XX
PN WO200015802-A1.
XX
PD 23-MAR-2000.
XX
PF 14-SEP-1999; 99WO-EP06781.
XX
PR 14-SEP-1998; 98GB-0020002.
XX
PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Ruelle J;
XX
DR WPI. 2000-271440/23.
XX
PT N-PSDB; AAA10700.
PT Novel BASB034 polynucleotides and polypeptides from Moraxella
catarrhalis used to prepare vaccines against bacterial infections

XX	Claim 3; Fig 2; 106pp; English.
PS	

This sequence represents a Moraxella catarrhalis BASB034 polypeptide from strain Mc2931 (ATCC 43617). The invention relates to BASB034 polypeptides from M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034 polynucleotides and polypeptides may be employed as research reagents and material for the discovery of treatments and diagnostics for diseases, particularly human diseases. They are particularly used to diagnose and treat M. catarrhalis infections. They can be used for diagnosis of disease, staging of disease, or determining response of an infectious organism to drugs. The polynucleotides may be used as a source for hybridization probes, and for screening of genetic mutations, serotype, organism or strain identification, identification of mutations in BASB034 sequences, and as components of arrays which are useful for diagnostic and prognostic purposes. The polypeptides can be used to produce antibodies. The polypeptides can also be used in vaccine formulations, and to identify agonists and antagonists. The polypeptides/antibodies, agonists and antagonists (which are bacteriostatic) are used for the treatment and prevention of diseases such as otitis media in infants and children, pneumonia in elderly, sinusitis, nosocomial infections and invasive diseases, and chronic otitis media with hearing loss. The polypeptides, agonists and antagonists are also used for screening of antibacterial drugs. The BASB034 products of the invention can be used screen for new antibacterial compounds that may target resistant bacteria.

SQ Sequence 442 AA;

Query Match	100.0%	Score 2360	DB 21	Length 442
Best Local Similarity	100.0%	Pred. 4.5e-219		
Matches 442	Conservative	0	Mismatches 0	Indels 0
				Gaps 0

Qy	1	mkvvsstltlslscfalfalldqoaaqvrnpafvdevrsemdagondelrldvqsatqsa	60
Db	1	mvvsstlslslslscfalfalldqqaqvnpvafvdevrsemdlqgdndelrldvqsaaqsa	60
Qy	61	smvdrnpdldehperelyttalekntklmgcsalndimklacudrlyvngetrpaulyktkrsi	120
Db	61	sdctanpldehperelyttalenkmlncsalngdmlmlaeudclvnhgetrpaulyktkrsi	120
Qy	121	rldefttwotlckrpoxytqettdprflmgnekmltkkdaqlaakotprslsstdld	180
Db	121	rldeftwglkxkprvlytqettdprflmgnekmltkkdaqlaeyaaqkftrslsstdld	180
Qy	181	rnntprlmsrripnpyulprfmgkgrnrsrptprsheakofprnfebradelkfqvsyvka	240
Db	181	rnntprlmsrripnpyulprfmgkgrnrsrptprsheakqftrnefrrelkfkqvsyvka	240
Qy	241	aedlmcgtdsdlmfgytqoshmqlfengksrpfvandyopelrltqpyvSDLpmdgkvrm	300
Db	241	aedlwgtdsdlmfgytqgsbwdlfngksrpfvandyopelrlcpvysdlwpgkvrm	300
Qy	301	gmgavnhnsgeaklsrbsmnraylmaagmeknltlmpmrigrktrkessgspodnprild	360
Db	301	gmavnhnsgeaklsrbsmnraylmaagmeknltlmpmrigrktrkessgsppodnprild	360
Qy	361	yygcgvdrlylodekksnlsgrvnrnpnsggaloldvvylylgrksgisgfoifogqosl	420
Db	361	yyvgvdrlylodekksnlsgrvnrnpnsggaloldvvylylgrksgisgfgifogqysl	420
Qy	421	ldynheatsfgvglmldnmgl	442
Db	421	ldynheatsfgvglmldnmgl	442

RESULT 2
AAV85271
ID AAV85271 standard; Protein; 442 AA
XX
AC AAV85271;
XX

DT 29-JUN-2000 (first entry)

DE BASB034 amino acid sequence #4.

KM Moraxella catarrhalis infection; BMSB034; diagnosis; straining;
 KM vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
 KM sinusitis; nosocomial infection; invasive disease; chronic otitis media;
 KM hearing loss; antibacterial drug

OS. *Moraxella catarrhalis*.

PN W0200015802-A1

PD 23-MAR-2000.

PF 14-SEP-1999; 99WO-EP06781.

PR 14-SEP-1998; 98GB-0020002.

PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS

PI Ruelle J;

DR WPI; 2000-271440/23

DR N-PSDB; AAA10703

Novel BASB034 polynucleotides and polypeptides from *Moraxella*

PT catarrhals used to prepare vaccines against bacterial infections

PS Claim 3; Page 69; 106pp; English.
xx

This sequence represents a *Moraxella catarrhalis* BASB034 polypeptide from strain Mc2969. The invention relates to BASB034 polypeptides from *M. catarrhalis* strains Mc2931, Mc2308, Mc2913 and Mc2965. The BASB034 polynucleotides and polypeptides may be employed as research reagents and material for the discovery of treatments and diagnostics for diseases, particularly human diseases. They are particularly used to diagnose and treat *M. catarrhalis* infections. They can be used for diagnosis of disease, staging of disease, or determining response of an infectious organism to drugs. The polynucleotides may be used as a source for hybridisation probes, and for screening of genetic mutations, serotype, organism or strain identification, identification of mutations in BASB033 sequences, and as components of arrays which are useful for diagnostic and prognostic purposes. The polypeptides can be used to produce antibodies. The polypeptides can also be used in vaccine formulations, and to identify agonists and antagonists. The polypeptides, antibodies, agonists and antagonists (which are bacteriostatic) are used for the treatment and prevention of diseases such as otitis media in infants and children, pneumonia in elderlies, sinusitis, nosocomial infections and invasive diseases, and chronic otitis media with hearing loss. The polypeptides, agonists and antagonists are also used for screening of antibacterial drugs. The BASB034 products of the invention can be used screen for new antibacterial compounds that may target resistant bacteria.

Sequence 442 AA;

Query Match	99.7%	Score 2354	DB 21	Length 442
Best Local Similarity	99.5%	Pred. No. 1.7e-218		
Matches 440; Conservative	1	Mismatches	1	Indels 0; Gaps 0

QY 1 MKSLSLTLFLSLSCFALIALIQQAQAVNPPAFVDERVSENDLGODNELPIDQSAQSA 60

Db 1 MKSLSLTLFLSLSPCFALIALQQAQAVNPPAFVDERVSENDIGQDNEPLDQSAQSA 60

QY 61 SDTAPALDHEBELYTTALEKKTKMLINCSALNDIMLACYDTFLVNGEPPAVIKTKRSI 120

Db 61 sttalanpdihebelyltalenkmlincsalngdlmlyacydcltvlngceppaviktkrsi 120

QY 121 RLDETIMQIKGPPQVIYVETDPIFLMGNGEKMGLTKDKAAQLEPYAAKQFPPLSLSDLD 180

Db 121 rldetlmwcltqkppqvvygvetcpdlflmgnekgnltkkdkaqlgyeakqkfcpplslstdld 180

QY 181 RNNTPLMSSRPNNPMYVLPDIFMHGKPNRSPNTPSHEAKOFTPNNEPAPELKFOVSVKVA 240
DB 181 rnnctpLwssrphnmpmyvlpdIfmhgkpnrsnpncpsneakqfcpneitrapelktfvsykvxa 240
QY 241 AEDLMGTSDDLMEGYTQOSHMOIFNGKNSRPFVHDYQPEIFLTQPVYSDLPMDGKVRMI 300
DB 241 aedlwgtddldwfytlqgshwqifngknsrpfvhdypceifltqpyvSDLpmdgkvrmI 300
QY 301 GNGAVHNSGESAKLSRSNNRAYLMAGMEKNLTWPRIMGRIFKRGSSQPDNDPDIID 360
DB 301 gngavhnsgeesaklsrswraylmgamewknltwprlmgrlfrkgsqsgpddndpdiId 360
QY 361 YGYGDVRFPLYOLENKSNTSGTVRNPFRSGKALQLDYVYPLGKISGYFOIFOGYGOSL 420
DB 361 ygygdvrflylqenksntsgtvrynprsgkaldyvyplgklsygfqlfgyygsal 420
QY 421 IDYNHEATSFVGLMLNDMWGL 442
DB 421 idynheatsfygvlmndwmgl 442

RESULT 3
AA85270 standard; Protein: 442 AA.
AY85270:
29-JUN-2000 (first entry)
BASB034 amino acid sequence #3.
Moraxella catarrhalis infection; BASB034; diagnosis; staging;
vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
sinusitis; nosocomial infection; invasive disease; chronic otitis media;
hearing loss; antibacterial drug.
Moraxella catarrhalis.
WO200015802-A1.
23-MAR-2000.
14-SEP-1999; 99WO-EP06781.
14-SEP-1998; 98GB-0020002.
(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
Ruelle J;
WPI: 2000-271440/23.
N-PSDB: AAA10702.
Novel BASB034 polynucleotides and polypeptides from Moraxella
catarrhalis used to prepare vaccines against bacterial infections
Claim 3; Page 68; 106pp; English.

This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
strain Mc2913. The invention relates to BASB034 polypeptides from
M. catarrhalis strains Mc2913, Mc2908, Mc2913 and Mc2969. The BASB034
polynucleotides and polypeptides may be employed as research reagents and
material for the discovery of treatments and diagnostics for diseases,
particularly human diseases. They are particularly used to diagnose and
treat M. catarrhalis infections. They can be used for diagnosis of
disease, staging of disease, or determining response of an infectious
organism to drugs. The polynucleotides may be used as a source for
hybridization probes, and for screening of genetic mutations, serotype,
organism or strain identification. Identification of mutations in BASB034
sequences, and as components of arrays which are useful for diagnostic
and prognostic purposes. The polypeptides can be used to produce
antibodies. The polypeptides can also be used in vaccine formulations,

CC and to identify agonists and antagonists. The polypeptides, antibodies,
CC agonists and antagonists (which are bacteriostatic) are used for the
CC treatment and prevention of diseases such as otitis media in infants and
CC children, pneumonia in eldersies, sinusitis, nosocomial infections and
CC invasive diseases, and chronic otitis media with hearing loss. The
CC polypeptides, agonists and antagonists are also used for screening of
CC antibacterial drugs. The BASB034 products of the invention can be used
CC screen for new antibacterial compounds that may target resistant
CC bacteria.
XX
SQ Sequence 442 AA:

Query Match 99.7%; Score 2352; DB 21; Length 442;
Best Local Similarity 99.3%; Pred. No. 2,7e-218;
Matches 439; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVSLSTLTLSTLSCEPAIIAQQAQAVPNPVAFVDEVRSENDIGDNELPIDVQSAQSA 60
DB 1 mkvslstltsltscefallaigqakavnpvatfdevrsendlgdnelpidvgsatgsa 60
QY 61 STDTPANPLDEHEPELYTTALENTKMLINCSALNQDIMRLACYTTLVHGEPNAYIKRKRSI 120
DB 61 stdtanpldehepelytalenktmlinccsalnqdimrlacydtllvhgepnayikrksI 120
QY 121 RUDETIWOTIKGKROYIYQETTDPIFLMGNEKGLRKKAQKOLEVAKOFTPLSLFDDLD 180
DB 121 rldetiwtikgkroyiyqettdpiflmgnekglrkkaqkolevakoftplslsfdld 180
QY 181 RNNTPLMSSRPNNPMYVLPDIFMHGKPNRSPNTPSHEAKOFTPNNEPAPELKFOVSVKVA 240
DB 181 rnnctpLwssrphnmpmyvlpdIfmhgkpnrsnpncpsneakqfcpneitrapelktfvsykvxa 240
QY 241 AEDLMGTSDDLMEGYTQOSHMOIFNGKNSRPFVHDYQPEIFLTQPVYSDLPMDGKVRMI 300
DB 241 aedlwgtddldwfytlqgshwqifngknsrpfvhdypceifltqpyvSDLpmdgkvrmI 300
QY 301 GNGAVHNSGESAKLSRSNNRAYLMAGMEKNLTWPRIMGRIFKRGSSQPDNDPDIID 360
DB 301 gngavhnsgeesaklsrswraylmgamewknltwprlmgrlfrkgsqsgpddndpdiId 360
QY 361 YGYGDVRFPLYOLENKSNTSGTVRNPFRSGKALQLDYVYPLGKISGYFOIFOGYGOSL 420
DB 361 ygygdvrflylqenksntsgtvrynprsgkaldyvyplgklsygfqlfgyygsal 420
QY 421 IDYNHEATSFVGLMLNDMWGL 442
DB 421 idynheatsfygvlmndwmgl 442

RESULT 4
AA85269 standard; Protein: 442 AA.
AY85269:
29-JUN-2000 (first entry)
BASB034 amino acid sequence #2.
Moraxella catarrhalis infection; BASB034; diagnosis; staging;
vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
sinusitis; nosocomial infection; invasive disease; chronic otitis media;
hearing loss; antibacterial drug.
Moraxella catarrhalis.
WO200015802-A1.
23-MAR-2000.
14-SEP-1999; 99WO-EP06781.

PR 14-SEP-1998; 98GB-0020002.
XX
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Ruelle J;
XX
DR WPI, 2000-271440/23.
DR N-PSDB; AAA10701.
XX
PT Novel BASB034 polynucleotides and polypeptides from Moraxella
PT catarrhalis used to prepare vaccines against bacterial infections -
XX
PS Claim 3; Page 67; 106pp; English.
XX
CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC strain Mc2908. The invention relates to BASB034 polypeptides from
CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
CC polynucleotides and polypeptides may be employed as research reagents and
CC material for the discovery of treatments and diagnostics for diseases,
CC particularly human diseases. They are particularly used to diagnose and
CC treat M. catarrhalis infections. They can be used for diagnosis of
CC disease, staging of disease, or determining response of an infectious
CC organism to drugs. The polynucleotides may be used as a source for
CC hybridization probes, and for screening of genetic mutations, serotype,
CC organism or strain identification, identification of mutations in BASB034
CC sequences, and as components of arrays which are useful for diagnostic
CC and prognostic purposes. The polypeptides can be used to produce
CC antibodies. The polypeptides can also be used in vaccine formulations,
CC and to identify agonists and antagonists. The polypeptides, antibodies,
CC agonists and antagonists (which are bacteriostatic) are used for the
CC treatment and prevention of diseases such as otitis media in infants and
CC children, pneumonia in elderly, sinusitis, nosocomial infections and
CC invasive diseases, and chronic otitis media with hearing loss. The
CC polypeptides, agonists and antagonists are also used for screening of
CC antibacterial drugs. The BASB034 products of the invention can be used
CC screen for new antibacterial compounds that may target resistant
CC bacteria.
XX
XX
SQ Sequence 442 AA;
Query Match 98.7%; Score 2330; DB 21; Length 442;
Best Local Similarity 98.6%; Pred. No. 3.5e-216;
Matches 436; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 MKVSLTLLTILSCFALIAIQAAVNPVAFVDEVSENDLGDNDLPIDVGSATOSA 60
Db 1 mkvslstlltlpofalialqgqavnpvafdevrskndlgdndlllgvgsatgsa 60
QY 61 STDTPANLDEHEPELYTTALENKTMLNCSALNODIMRLACYDTLVHGETPAVITKRSI 120
Db 61 stdtpanldehepeelyttalenktmlncsalnodimrlacydtlvhgetpaviktrsi 120
QY 121 RLDEFINOTIGKPOVYTOETDPTFLMNGKMTKRDAQOLEAAQOFPLSTSPFLD 180
Db 121 rldefinotigkpoivytoetdptflmngkmtkrdaqoleaaqofplstspfld 180
QY 121 rldelilwtllygkpvvygetcdpflimgnekmltkkdaqlvyaakqfiprlisifld 180
Db 181 RNNTPILMSRPHNPYVPIFMHGKPNRSPTPSHEAKQFTPNERABELKFOVSVKKA 240
QY 181 rnntpilmsrphnpyvpifmhgkpnrsptpsheakqftpnerabelkfovsvkka 240
Db 181 rnntrplwsrphnpyvpiifmhgkpnrsptpsheaqfipnetrapelkfqvsvkka 240
QY 241 AEDLMGTDSDLMFGYTOQSHQOIFENGKNSRPRVHDYQPEIFLQPVYSDLPMDGKVNMT 300
Db 241 aedlmgtddslwfytgqshwqifngknsrprvhdypelfltqpvysdlpmdgkvmt 300
QY 301 GMSGAVHNSGSASAKTSRSMNRAVYLMAGMEKNLTVMPTMGRIEFGSGSQPDNDPDLID 360
Db 301 gmsgavhnsngsaksrswmncraylmgamewknlvmpmrltwtglfkgsgsqgddndpdlid 360
QY 361 YYGGVDFVFLQLENNKNSISGTVRYNPRSGKALQLDVYVPLGKISGFFQFGYGSOL 420
Db 361 yygvdvfllyqlenknlsigtrvynprsgkalqldvyvplgklsyfgqlfgyggsal 420

QY 421 IDYNHEATSEGVGLMNDMGL 442
Db 421 idynheatsfgvglmndmgl 442
RESULT 5
AAV75156
ID AAV75156 standard; Protein: 370 AA.
XX
AC AAV75156;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria gonorrhoeae ORF 582 protein sequence SEQ ID NO:1786.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
KW antibacterial; gene therapy.
XX
OS Neisseria gonorrhoeae.
PN
PN W0957280-A2.
PD 11-NOV-1999.
XX
XX 30-APR-1999; 99WO-US09346.
PE
XX 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103796.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
XX (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
XX Fraser C, Galeotti C, Grandi G, Hickey E, Malignani V, Mora M;
PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
XX WPI: 2000-062150/05.
DR N-PSDB; AA253918.
XX
PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics -
XX
PS Claim 2; Page 903; 1453pp; English.
XX
XX AA253015 to AA254536, AA254577 to AA254615, and AAV74253 to AAV75941
CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
CC and polypeptides. AA254537 to AA254576 and AA254616 to AA254473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicemia) to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
SQ Sequence 370 AA;
Query Match 34.8%; Score 821; DB 21; Length 370;
Best Local Similarity 44.4%; Pred. No. 1.2e-70;
Matches 162; Conservative 69; Mismatches 110; Indels 24; Gaps 8;
QY 87 INCSALNODIMRLACYDTLVHGETPAVI-----KTKRSIRLDETWTOTI-KGKPOVITOE 140

CC sequence of H. pylori (ATCC 55679) was determined from overlapping
CC contigs generated by mechanically shearing the bacterial DNA. The
CC sequences were analysed for ORF or at least 180 nucleotides, and the
CC predicted coding regions defined by computer evaluation. To identify
CC likely H. pylori antigens for vaccine development, the amino acid
CC sequences predicted from various ORF were analysed for significant
CC homology to other known or exported membrane proteins. Having identified
CC and determined the sequences of interest, particular regions can be
CC isolated from H. pylori by PCR amplification for recombinant polypeptide
CC production, e.g. in E. coli hosts.

CC
XX
SQ Sequence 253 AA:

Query Match 6.6%; Score 156.5; DB 18; Length 253;
Best Local Similarity 25.6%; Pred. No. 9.6e-07;
Matches 56; Conservative 26; Mismatches 72; Indels 65; Gaps 8;

QY 157 KKDAKOLEYAAKQFTPLSFDLDRNNTPLMSSRPNNPMYVLPFMHGKNSRPNPSHE 216
DB 72 KkYlmmddYlgYfLpFhYsf-----rPlfGwYhPnlP----- 105
QY 217 AKQFTNERAPRLKQVSVKKAEDLMGTDSDLMFGYTOQSHWQIFNGKNSRPFKVDH 276
DB 106 ---Yqrnef---kfqsftrypvfrhllwtkgtlYlayrqtWfqlYndpgsqpmmin 157
QY 277 YQPEITLQPVYSDLPMDCGV---RMIGCAVHHSNG-ESAKLSRSNNRAYLMAGMEKN 332
DB 158 fmpellyvYpI-nfkpfqgkIgfseIwlgwqhlshngvgagcyqphn----- 204
QY 333 LTVMPRIWGRIFKEGSGSQ--PDDNPDIIDYGYGVDYRF 369
DB 205 -----kegnpeqfpgqpvIvXkdYngqkdvW 231

RESULT 13
AAW70991
ID AAW70991 standard; Protein; 1686 AA.

XX
AC AAW70991;

XX
DT 19-OCT-1998 (first entry)

XX
DE Human class II PI3 kinase-C2alpha.

XX
KW Human; class II phosphoinositide lipid kinase; PI3 kinase;
KW PIK3-C2alpha; class II; resistance; Wortmannin; LY294002.

XX
OS Homo sapiens.

XX
PN WO9832864-A2.

XX
PD 30-JUL-1998.

XX
PF 27-JAN-1998; 98MO-GB00244.

XX
PR 28-JAN-1997; 97GB-0001652.

XX
PA (LUDW-) LUDWIG INST CANCER RES.

XX
PI Domln J, Waterfield MD;

XX
DR WPI; 1998-427960/36.

XX
DR N-PSDB; AAV42920.

XX
PT New nucleic acid encoding phospho:inositol kinase 3-C2 alpha or its
PT fragments - useful for, e.g. treatment of tumour cells where
PT phenotype is associated with expression of kinase

XX
PS Claim 3; Fig 1; 52pp; English.

XX
CC The present sequence represents a human class II phosphoinositide lipid
CC (PI3) kinase designated PIK3-C2alpha. It is characterised as a class II

CC kinase due to the presence of a conserved C2 domain found in murine and
CC Drosophila class II PI3 kinases, its apparent lack of a p85 binding site
CC and a substrate affinity to inositol lipids PtdIns and PtdIns(4)P. The
CC protein has resistance to PI3 kinase inhibitors Wortmannin and LY294002.
CC Antibodies against the protein (optionally humanised), are used to
CC identify class II PI3 kinases. Antisense sequences, antibodies or
CC dominant negative mutants of the PI3-C2alpha protein, are useful in human
CC or veterinary medicine to block class II kinases. They can be used to
CC treat tumour cells where the phenotype is associated with expression of
CC PI3-C2alpha protein.

CC
XX
SQ Sequence 1686 AA:

Query Match 5.0%; Score 117; DB 19; Length 1686;
Best Local Similarity 19.2%; Pred. No. 0.13;
Matches 86; Conservative 67; Mismatches 154; Indels 140; Gaps 21;

QY 45 QDNELPIDY-----QSATQASITDTAMPDLE-----HEPELYTTALENKMMLIN----- 88
DB 522 eddepydlnkhlyqIekpkeamttrpveellidsyngvelalqlenqtravdqvlav 581
QY 89 ---CSALNODIMRLACYDTLVHGETTPAVIKTKRSIRIDETIMQTIKGPQVIYQETDPI 145
DB 582 rklcsald-gvelal-----tesvklkkravnlpr-----ktadv 618
QY 146 FLMGNEKGMILTKDAKOLEYAAKQFTPLSFDLDRNNTPLMSSRPNNPMYV-----LP 199
DB 619 sltge-----dtsrst-rgslnpenpYvgslnqlta 651
QY 200 IF-----HHGPNRSPPNPSHAKQFTNERAPRLKQVSVKKAEDL---WGTSDLM 252
DB 652 Iydlrlrhansgrspdcassksvkaewtltelqtlf---aaglnsswnsnjky 707
QY 253 FGYYTOQSHWQIFNGKN-SRPF---RVHDYQPELFLTO-----PV-YSDLPMWGGKVRMI 300
DB 708 YllcsIsh---ngkdlfkplqekkvgtYknffYllkwdeIlfplqlsqplsvlht 763
QY 301 GMGAVHHSNGESAKLSRSNN-----RAYLMAGMEKNLTVMPRIWGRIFK 345
DB 764 lfglngssgsspsdnkqkYqpealqkvsIplcdfrrflcY-----tkllylw----- 812
QY 346 EGSQSGPDDNPDIIDYGYGVDYRFYOLENKSNTSGVRRYNNRSGCALDLDYVPLGK 405
DB 813 --tsshtnsvpyvtckkygYmerIvqvdFpsafdllytppqvdrlsqhmlcIend 870
QY 406 ISGYFOIFQGYGOSLIDYNNHEATSGV 432
DB 871 lky-----klldllhkdsIgl 887

RESULT 14
AAP70420
ID AAP70420 standard; protein; 800 AA.

XX
AC AAP70420;

XX
DT 20-JAN-1991 (first entry)

XX
DE Sequence encoded by cellulase gene derived from Bacillus sp. No. 1139.

XX
KW Enzyme; cellotriose; cellotetraose; hydrolysis.

XX
OS Bacillus sp. No. 1139.

XX
FH Key Location/Qualifiers

XX
FT Peptide 1..30

XX
FT Protein 31..800

XX
PN JP62232386-A.

XX
PD 12-OCT-1987.

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:14:03 ; Search time 36.79 Seconds
(without alignments)
270.358 Million cell updates/sec

Title: US-09-787-083-2
Perfect score: 2360
Sequence: 1 MKVSLSTLTLTSLSCFALIA.....YNHRTSRGVGLMINDMGL 442

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues.

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101.5	4.3	1726	2	US-08-609-049A-30
2	101.5	4.3	1726	4	US-09-170-996-30
3	100.5	4.3	1658	2	US-08-609-049A-13
4	100.5	4.3	1658	4	US-09-170-996-13
5	95.5	4.0	682	3	US-08-481-435-6
6	91.5	3.9	1024	4	US-09-091-117-5
7	89.5	3.8	503	1	US-07-946-497-2
8	89.5	3.8	503	1	US-08-483-322-2
9	89.5	3.8	503	2	US-08-478-882-2
10	89.5	3.8	857	1	US-07-717-331F-2
11	87	3.7	537	2	US-08-633-879C-2
12	87	3.7	985	5	PCR-US96-03916-6
13	87	3.7	985	5	PCR-US96-03916-66
14	86.5	3.7	666	4	US-08-961-083-2
15	86	3.6	781	1	US-08-373-134D-2
16	86	3.6	781	2	US-09-114-637-2
17	86	3.6	1627	1	US-07-665-792E-9
18	84	3.6	522	6	RE34605-6
19	84	3.6	816	1	US-07-731-157A-4
20	84	3.6	816	1	US-08-229-444B-2
21	84	3.6	816	2	US-08-541-780-4
22	83.5	3.5	774	3	US-08-802-632-2
23	83.5	3.5	774	3	US-09-073-354-1
24	83.5	3.5	774	3	US-08-656-005A-1
25	83.5	3.5	774	4	US-09-073-259-1
26	83.5	3.5	774	4	US-09-363-095-1
27	83.5	3.5	774	4	US-09-418-027-1

28	83	3.5	355	1	US-07-946-497-5	Sequence 5, App1
29	83	3.5	355	1	US-08-483-322-5	Sequence 5, App1
30	83	3.5	355	2	US-08-478-082-5	Sequence 5, App1
31	82.5	3.5	535	2	US-08-564-972-1	Sequence 1, App1
32	82.5	3.5	535	4	US-09-171-969-9	Sequence 9, App1
33	82.5	3.5	657	4	US-09-306-593-2	Sequence 2, App1
34	82.5	3.5	663	1	US-08-441-139-7	Sequence 7, App1
35	82	3.5	334	2	US-08-359-850-4	Sequence 4, App1
36	82	3.5	527	2	US-08-592-126-145	Sequence 145, App
37	82	3.5	527	2	US-08-687-080-48	Sequence 48, App1
38	82	3.5	535	2	US-08-633-879C-4	Sequence 4, App1
39	82	3.5	797	4	US-09-086-912-2	Sequence 2, App1
40	82	3.5	1022	3	US-08-772-270A-2	Sequence 2, App1
41	82	3.5	1088	3	US-08-633-768A-1	Sequence 1, App1
42	81	3.4	320	2	US-08-245-511-4	Sequence 4, App1
43	81	3.4	320	2	US-08-600-993A-4	Sequence 4, App1
44	80.5	3.4	354	1	US-07-946-497-4	Sequence 4, App1
45	80.5	3.4	354	1	US-08-483-322-4	Sequence 4, App1

ALIGNMENTS

RESULT 1
US-08-609-049A-30
Sequence 30, Application US/08609049A
Patent No. 5948664
GENERAL INFORMATION:
APPLICANT: Williams, Lewis T.
APPLICANT: Molz, Lisa
TITLE OF INVENTION: No. 5948664e1 PI 3-Kinase Polypeptides
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/609,049A
FILING DATE: 29-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-063700US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 1726 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-609-049A-30

Query Match 4.3% Score 101.5; DB 2; Length 1726;
Best Local Similarity 18.6% Pred. No 0.47;
Matches 98; Conservative 80; Mismatches 189; Indels 161; Gaps 21;
QY 1 MKVSLSTLTLTSLSCFALIAIOQAQVNPVAFV-----EYRSNDLQGD 46
DB 465 VKVSLIEGLQLPYTFCDVSVTEVIEIIMQALSVHVDLQNVGVSYLTKVCGGEVLYN 524

```
OY 47 NELPIDVOSATQASTDTANPLDEHEPELYTTALENKTMLNCSALNODIMRLACYDTLV 106
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 525 NHCLSGHEHIONCRKMDT-----EIKLQLTLTSLAMCONLTARTADD--- 565
OY 107 HGPTRAVIKTRSRITLDETIMOTIKGKPOVITYOETTDPIF----- 146
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 566 --EAP-----VDLNKLYLQIEKPYKEVMTNRHPEVELDSYHYVELALQTEHONRAV 615
OY 147 -----LMGNEKGMLTKKDAKOLEYAA-----KOPTPLSLFDDLRNNTPLMS 188
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 616 DQVIAVKRKCISALDGVETPSYTEA-VKKIKRAVNLPRKNSADYTSLSGSDTRKSTGSG 674
OY 189 SRPHNPMYVLP-----IFMHGKPNR-----SPNTPSHEAKOPTNEFRAPELKFOV 234
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 675 LNPEHPVQVSMOHLTRTRYDLRLHANSSRCSTGCCPRGSRNKKEAMTATE-----OL 726
OY 235 SVKYVAEDL---WGTDSDLMGVYTOOSHMOIFNGKN-SRPF---RVHDOPEIFLTO-- 285
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 727 QFTYVAHGISNMVSNKEKYLICLSLH---NGKDLFKPIQSKKVGTYKNFEYLKMD 782
OY 286 ----PV-YSDLPMGKYRMIGMGAHVHSHNGESAKLSRSWN-----RAVL 324
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 783 ELIIFPIQSLPLESVLHLLFGLVNLQSSGSSPPSNKQKPEALGKVSLLTLPFKRFL 842
OY 325 MAGMEMKMLTYMPRIWGRIFKEGSGSQPDNDPDIIDYGYGVRFLYOLENKSNTGTVR 384
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 843 TCG-----TKLLYLW-----TSSHTNSIPGALPKKSYMERIVLQVDFPSPAFDIY 889
OY 385 YNPRSGKALQDLYVPLGKIGISGYFOIFGQGSGLIDYHNHATSEGV 432
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 890 TSPQIDRNITQODKLETLESDIKG-----KLDIITHDSSFG 927
```

RESULT 2

```
US-09-170-996-30
; Sequence 30, Application US/09170996
; Patent No. 6291220
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Molz, Lisa
; APPLICANT: Chen, Yen-Wen
; TITLE OF INVENTION: No. 6291220e1 PI 3-kinase Polypeptides
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/170.996
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/609,049
; FILING DATE: 29-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-063700US
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1726 amino acids
; TYPE: amino acid
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-170-996-30
```

Query Match 4.3%; Score 101.5; DB 4; Length 1726;
Best Local Similarity 18.6%; Pred. No. 0.47;
Matches 98; Conservative 80; Mismatches 189; Indels 161; Gaps 21;

```
OY 1 MKVSLS-----TLTSLISCFALIAIQAAQVNPVAFVD-----EVRSENDLQD 46
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 465 VAVSIEIEGLQLPYFTCDVSVTEIIMQALSWHDLNQVDGSIYIKCGQGEVLCN 524
OY 47 NELPIDVOSATQASTDTANPLDEHEPELYTTALENKTMLNCSALNODIMRLACYDTLV 106
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 525 NHCLSGHEHIONCRKMDT-----EIKLQLTLTSLAMCONLTARTADD--- 565
OY 107 HGPTRAVIKTRSRITLDETIMOTIKGKPOVITYOETTDPIF----- 146
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 566 --EAP-----VDLNKLYLQIEKPYKEVMTNRHPEVELDSYHYVELALQTEHONRAV 615
OY 147 -----LMGNEKGMLTKKDAKOLEYAA-----KOPTPLSLFDDLRNNTPLMS 188
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 616 DQVIAVKRKCISALDGVETPSYTEA-VKKIKRAVNLPRKNSADYTSLSGSDTRKSTGSG 674
OY 189 SRPHNPMYVLP-----IFMHGKPNR-----SPNTPSHEAKOPTNEFRAPELKFOV 234
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 675 LNPEHPVQVSMOHLTRTRYDLRLHANSSRCSTGCCPRGSRNKKEAMTATE-----OL 726
OY 235 SVKYVAEDL---WGTDSDLMGVYTOOSHMOIFNGKN-SRPF---RVHDOPEIFLTO-- 285
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 727 QFTYVAHGISNMVSNKEKYLICLSLH---NGKDLFKPIQSKKVGTYKNFEYLKMD 782
OY 286 ----PV-YSDLPMGKYRMIGMGAHVHSHNGESAKLSRSWN-----RAVL 324
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 783 ELIIFPIQSLPLESVLHLLFGLVNLQSSGSSPPSNKQKPEALGKVSLLTLPFKRFL 842
OY 325 MAGMEMKMLTYMPRIWGRIFKEGSGSQPDNDPDIIDYGYGVRFLYOLENKSNTGTVR 384
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 843 TCG-----TKLLYLW-----TSSHTNSIPGALPKKSYMERIVLQVDFPSPAFDIY 889
OY 385 YNPRSGKALQDLYVPLGKIGISGYFOIFGQGSGLIDYHNHATSEGV 432
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 890 TSPQIDRNITQODKLETLESDIKG-----KLDIITHDSSFG 927
```

RESULT 3

```
US-08-609-049A-13
; Sequence 13, Application US/08609049A
; Patent No. 5948664
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Molz, Lisa
; APPLICANT: Chen, Yen-Wen
; TITLE OF INVENTION: No. 5948664e1 PI 3-kinase Polypeptides
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,049A
; FILING DATE: 29-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
```


NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-063700US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1658 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-609-049A-13

Query Match 4.3%; Score 100.5; DB 2; Length 1658;
Best Local Similarity 18.9%; Pred. No. 0.56;
Matches 101; Conservative 79; Mismatches 194; Indels 159; Gaps 23;

QY 5 LSTLTLSILSCFALLAIQ---QAQAVNPVAFVDEVRS-----ENDLGODNE 48
DB 381 LSPVTYGRNMGCBENASVKSIEIEGLDLPVTFCDVSVTEIIMQALCWVHDLNQ--- 437
QY 49 LPIDVOS-----ATQASSTDFANPLDEHE--PELYTTALENKTMILNCSALNDIMRLAC 101
DB 438 --VDVGSYILKVCQGEVLIQNNHCLSGSHEHIONCRKWDTEIKQLTLISAMCONLARTAE 495
QY 102 YDPLVHGEPRAVIKTRKIRLDETIMOTIKGPQVIYQETTPDF----- 146
DB 496 DD-----EAP-----VDLANKYLIQIEKPYKEVMIHRHVEELDSYHYVELALOTEN 542
QY 147 -----LMGNEKMLTKKDAQOLEYAA-----KQPTPLSLSFDLDRNN 183
DB 543 QHRAVDQVIAKVRKICSLDGVETPSTEA-VKKLKAVALPRKKSADVTSLSGSDTRKN 601
QY 184 TPLWSSRPHNPVYLP-----IFMKGPNR---SPNTPSHAKOFTPREFAPE 229
DB 602 STKSLPENPVOSMDHLTTATYDLIRLHANSRCSTGCPGRSRIKKEAWTATE----- 656
QY 230 LKFOVSVKVKAABDL---WGTDSDLMEGYTQOSHMOIFNGKN--SRPE---RVHDYOPEIF 282
DB 657 --QDFTVYAAHGISSNMVSNYEKYLICLSH---NCKDLFKPIQSKKVGTYKNKFY 709
QY 283 LTO-----PV-YSDLPMDGKVRMIGMGAHVHNSGSAKLSRSWN----- 320
DB 710 LKMWDELIIFPIQISQPLESVLHLTLFGVLNOSGSSPSDNKORKPEALGKVSLLTFD 769
QY 321 -RAYLMAGMEMKMLTYMPRIWGRIFKESGSGQPDNDILDYGYGVDVRLYOLENKSNI 379
DB 770 FKRFLLCG-----TKLLYLM-----TSSHTNSIPGALPKKSYMERIVLQVDFPSPA 816
QY 380 SGTVRYNPRSGKALQLDVYVPLGKISGYFOIFQGGSLIDYNHEATSGV 432
DB 817 FDIITYSPQIDRNIIQODKLETLESIDIKG-----KLDDIHRDSSFGL 859

RESULT 4
US-09-170-996-13
Sequence 13, Application US/09170996
Patent No. 6291220
GENERAL INFORMATION:
APPLICANT: Williams, Lewis T.
APPLICANT: Moliz, Lisa
APPLICANT: Chen, Yen-Wen
TITLE OF INVENTION: No. 6291220e1 PI 3-Kinase Polypeptides
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA

ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/170,996
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/609,049
FILING DATE: 29-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-063700US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1658 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-170-996-13

Query Match 4.3%; Score 100.5; DB 4; Length 1658;
Best Local Similarity 18.9%; Pred. No. 0.56;
Matches 101; Conservative 79; Mismatches 194; Indels 159; Gaps 23;

QY 5 LSTLTLSILSCFALLAIQ---QAQAVNPVAFVDEVRS-----ENDLGODNE 48
DB 381 LSPVTYGRNMGCBENASVKSIEIEGLDLPVTFCDVSVTEIIMQALCWVHDLNQ--- 437
QY 49 LPIDVOS-----ATQASSTDFANPLDEHE--PELYTTALENKTMILNCSALNDIMRLAC 101
DB 438 --VDVGSYILKVCQGEVLIQNNHCLSGSHEHIONCRKWDTEIKQLTLISAMCONLARTAE 495
QY 102 YDPLVHGEPRAVIKTRKIRLDETIMOTIKGPQVIYQETTPDF----- 146
DB 496 DD-----EAP-----VDLANKYLIQIEKPYKEVMIHRHVEELDSYHYVELALOTEN 542
QY 147 -----LMGNEKMLTKKDAQOLEYAA-----KQPTPLSLSFDLDRNN 183
DB 543 QHRAVDQVIAKVRKICSLDGVETPSTEA-VKKLKAVALPRKKSADVTSLSGSDTRKN 601
QY 184 TPLWSSRPHNPVYLP-----IFMKGPNR---SPNTPSHAKOFTPREFAPE 229
DB 602 STKSLPENPVOSMDHLTTATYDLIRLHANSRCSTGCPGRSRIKKEAWTATE----- 656
QY 230 LKFOVSVKVKAABDL---WGTDSDLMEGYTQOSHMOIFNGKN--SRPE---RVHDYOPEIF 282
DB 657 --QDFTVYAAHGISSNMVSNYEKYLICLSH---NCKDLFKPIQSKKVGTYKNKFY 709
QY 283 LTO-----PV-YSDLPMDGKVRMIGMGAHVHNSGSAKLSRSWN----- 320
DB 710 LKMWDELIIFPIQISQPLESVLHLTLFGVLNOSGSSPSDNKORKPEALGKVSLLTFD 769
QY 321 -RAYLMAGMEMKMLTYMPRIWGRIFKESGSGQPDNDILDYGYGVDVRLYOLENKSNI 379
DB 770 FKRFLLCG-----TKLLYLM-----TSSHTNSIPGALPKKSYMERIVLQVDFPSPA 816
QY 380 SGTVRYNPRSGKALQLDVYVPLGKISGYFOIFQGGSLIDYNHEATSGV 432
DB 817 FDIITYSPQIDRNIIQODKLETLESIDIKG-----KLDDIHRDSSFGL 859

RESULT 5
US-08-481-435-6

Sequence 6, Application US/08481435
Patent No. 6027906
GENERAL INFORMATION:
APPLICANT: Balgansh, Tanjore S
APPLICANT: Town, Christine
TITLE OF INVENTION: No. 6027906el Polypeptides
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: White & Case
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,435
FILING DATE: 10-JUL-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IN 580/MAS/94
FILING DATE: 01-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9404072-2
FILING DATE: 24-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Sterner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-151
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 682 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-481-435-6

Query Match 4.0%; Score 95.5; DB 3; Length 682;
Best Local Similarity 20.4%; Pred. No. 0.43;
Matches 92; Conservative 66; Mismatches 191; Indels 103; Gaps 24;

QY 13 LSCFAIQAQAVPNVAFVDEVRENDL-GQDNELPIDVQATQ-----AST 62
DB 5 LSESKIVATTSKIDNKQIADLGSERRVNAQANDIPTDLYKAIVSIEDHREFDHKGI 64
QY 63 DTANPLDHEPELYTTALENTMLINCALNODIMRLACYDTLVGNETPAVIKTRIRL 122
DB 65 DIIRLGLFLRLNLSQSG-----SALTQOLIKLYFTSTSDQITIS-----RK 110
QY 123 DETIWTI---GKRPQVYIYQETDPIFLMGNEKGLTK-----KDAKQLE----- 164
DB 111 AQEAMLAIQLEKATKQELIYIINKYVMSNGYMGQTAQNYKGLDNLNLSPLQALLA 170
QY 165 ---YAAKQFTPLS-LSFDDLNNTPLMSSRRPNPMTVLP-IFEMHGKPNNSPNTPSHEAKQF 220
DB 171 GMPQAPNQPDPYSHPEAQAODRRNLVL--SEMKNGYI-----SAEQYERAVNTPITDGLQS 224
QY 221 TPNERARE-----LKFOYS-VKVKAAEDLMGTSDIMWEGTYQ--QSH-WOLFNGKNSRP 271
DB 225 LKASNSNYAYMDNYLKEVINYVEETGTNLLTTGMDVYTNNDQEKHLMDIYNDVEYA 284
QY 272 FRVHDYQPEIFLTQPVYSDLPWDGKVRMIGKGAHVHSN-----GESAKLSRSNNRAYLM 325
DB 285 YPDDELQ-----VASTIYDVS-NGKV-IAQLGARHSSNVSGINQAVETNDW----- 331

QY 326 AGMEKKNLTF-VMPRIWGIIEKSGSDPDNDLIDYQ-----YGDVRFILYQ 372
DB 332 -GSTKKPTIDYAPALEYGVY-ESTATIVHDEP--YNYPGTNPVYMDRGYFGNITLOYA 387
QY 373 LKNSNIGSTVRYN-----PRSGKALQDLY 398
DB 388 LQOSRNPVAVETLKNKVGINRAKTFPLNGLGDY 419

RESULT 6

US-09-091-117-5
Sequence 5, Application US/09091117
Patent No. 6171589
GENERAL INFORMATION:
APPLICANT: The University of Melbourne
TITLE OF INVENTION: Mycoplasma Recombinant Polypeptides and
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: GREENLEE, WINNER and SULLIVAN P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: United States of America
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/091,117
FILING DATE: 12 JUNE 1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU96/00803
FILING DATE: 13-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PNT127
FILING DATE: 13-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: WINNER, Ellen P.
TELECOMMUNICATION INFORMATION:
TELEPHONE: +1 303 499 8080
TELEFAX: +1 303 499 8089
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1024 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Mycoplasma genitalium
US-09-091-117-5

Query Match 3.9%; Score 91.5; DB 4; Length 1024;
Best Local Similarity 20.2%; Pred. No. 2.2;
Matches 86; Conservative 61; Mismatches 172; Indels 107; Gaps 22;

QY 1 MKVS-LSITLTLSILSCFAIQAQAVPNVAFVDEVRENDLGDNDLPI-DVQATQ 58
DB 1 MKLSTITTCISISGAFGTAT-----ALFTYVALKLNHQQQNTKQON--PIKDIRGLIN 54
QY 59 SASSTANPLDHEPELYTTALENTMLINCALNODIMRLACYDTLVGNETPAVIKTR 118
DB 55 NQVQNTIPL--HQIVVEVT--NNKAIYDKAPQKFL-----AKSALNNKL 98
QY 119 SIRLDETIWTIKGKRPQVYIYQETDPIFLMGNEKGLTKDAKQLEVAKO-FTPLSLSF 177
DB 99 QVEFDFKFLRT-----GVINALNADLKEWIDOTLFIPOSEF 135
QY 178 DLDNRNTPLMSSRRPNPMTVLP-IFEMHGKPNRSP--NTPSHEAKQFTPNFRARPELKFOVS 235

Db 136 DLSANKLNLTLNSQSEVSLDLFEFTNFSDKNQPLKLPFGSVVYVANE-----SYTSYK 191
QY 236 VKKAADLMGTSDLMFGTQOOSHQIFNGKSRPRRVVDYQPEI-FLTQPYSDL-PW 293
Db 192 ATLQKAKVLYTSRADSHVGSYVAIPYVSLNGKYNDFSNFDPKSNINFAPKKNYVNALNPF 251
QY 294 DGKVRIGMGAVHNS--NGESAK-----LSRSMNRAVYLWAGHWNKLLVMPRIKGRIFK 345
Db 252 EAQOYLVGCGKFLNOKVKNADVKNDINNHIETQFNVAKITA-----TLGKAFK 300
QY 346 ---EGSGSQP-----DDNPILDY--GYGDV-----RFLYOL-- 373
Db 301 QFGEHNGQPLSLKVLKSLGSLNNEFKOLFNVVRGLDGFVSDLIQSSQSNKKTYYQLLF 360
QY 374 ENKSN 379
Db 361 ENKTTI 366

RESULT 7
US-07-946-497-2
; Sequence 2, Application US/07946497
; Patent No. 5506119

; GENERAL INFORMATION:
; APPLICANT: HERRLICH, Peter
; APPLICANT: GUNTHERT, Ursula
; APPLICANT: MATZKU, Siegfried
; APPLICANT: WENZL, Achim
; TITLE OF INVENTION: VARIANT CD4 SURFACE PROTEINS, DNA
; TITLE OF INVENTION: SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
; TITLE OF INVENTION: AS WELL AS THEIR USE IN DIAGNOSIS AND THERAPY
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/946,497
; FILING DATE: 19921109

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16915/145
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 503 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-946-497-2

Query Match 3.8%; Score 89.5; DB 1; Length 503;
Best Local Similarity 18.4%; Pred. No. 1.1;
Matches 57; Conservative 32; Mismatches 96; Indels 125; Gaps 14;

QY 40 ENLDGDNELPIDVQASATGDTANPLDHEPELYTALLENKTMILNCSALNODIMRL 99
Db 169 QEDIDASNIDEVSS-----GSTIEKSTPEGYILHTDPT---SQPTGGRD----- 212

QY 100 ACYDTLVHGENPAVYIKTKRSI-----RLDETLMQTIKGRPOVYQETDPIFLMGNEK 153
Db 213 ---DAFFIGSTLTIATTPWVSAHTKONQERTQNPILHNSPEVLQTT----- 258
QY 154 MLTKKDAKOLEYAKOFTPLPSFLDRLNNTPL-----WSSRPNMAYVLPIFMHKGKPRS 209
Db 259 -----RMT-----DIDRNSTSAHGEWVTOEPQPF----- 283
QY 210 PNTPSHEAKOFTPNEFPAPELKFOVSVKKAEDLW-----CTSDLMFGTQOS 259
Db 284 -NNHEYODEEETBH-----ATSTTWADPNSTTEAATQKEKF-----EN 322
QY 260 HMOIFNGKN-----SRFRVHDYQPEIFLQPYSDLPW-----DGKVRMTGMG 303
Db 323 EMQ---GKNPPSEDSHTVETGTTASAHNNHPSORMTQSOEDSWTDFDPISHPMGQG 379
QY 304 AVHNSGESA 313
Db 380 HOTESKCHSS 389

RESULT 8
US-08-483-322-2
; Sequence 2, Application US/08483322
; Patent No. 5760178

; GENERAL INFORMATION:
; APPLICANT: HERRLICH, Peter
; APPLICANT: GUNTHERT, Ursula
; APPLICANT: MATZKU, Siegfried
; APPLICANT: WENZL, Achim
; TITLE OF INVENTION: VARIANT CD4 SURFACE PROTEINS, DNA
; TITLE OF INVENTION: SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
; TITLE OF INVENTION: AS WELL AS THEIR USE IN DIAGNOSIS AND THERAPY
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,322
; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/946,497
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16915/145
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 503 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-483-322-2

Query Match 3.8%; Score 89.5; DB 1; Length 503;
Best Local Similarity 18.4%; Pred. No. 1.1;
Matches 57; Conservative 32; Mismatches 96; Indels 125; Gaps 14;

[illegible]

```

1      RESULT          9
2      US-08-478-882-2
3      Sequence 2, Application US/08478882
4      Patent No. 5885575
5      GENERAL INFORMATION:
6      APPLICANT: HERRLICH, Peter
7      APPLICANT: FONTA, Helmut
8      APPLICANT: GUENTHERT, Ursula
9      APPLICANT: MATZKU, Siegfried
10     APPLICANT: WENZL, Achim
11     TITLE OF INVENTION: VARIANT CD44 SURFACE PROTEINS, DNA
12     TITLE OF INVENTION: SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS
13     TITLE OF INVENTION: AS WELL AS THEIR USE IN DIAGNOSIS AND THERAPY
14     NUMBER OF SEQUENCES: 8
15     CORRESPONDENCE ADDRESS:
16     ADDRESSEE: Foley & Lardner
17     STREET: 3000 K Street, N.W., Suite 500
18     CITY: Washington, D.C.
19     COUNTRY: USA
20     ZIP: 20007-5109
21     COMPUTER READABLE FORM:
22     MEDIUM TYPE: Floppy disk
23     COMPUTER: IBM PC compatible
24     OPERATING SYSTEM: PC-DOS/MS-DOS
25     SOFTWARE: PatentIn Release #1.0, Version #1.25
26     CURRENT APPLICATION DATA:
27     APPLICATION NUMBER: US/08/478,882
28     FILING DATE:
29     CLASSIFICATION: 435
30     PRIOR APPLICATION DATA:
31     APPLICATION NUMBER: US/07/946,497
32     FILING DATE: 1992/11/09
33     ATTORNEY/AGENT INFORMATION:
34     NAME: BENT, Stephen A.
35     REGISTRATION NUMBER: 29,768
36     REFERENCE/DOCKET NUMBER: 16915/145
37     TELECOMMUNICATION INFORMATION:
38     TELEPHONE: (202)672-5300
39     TELEFAX: (202)672-5399
40     TELEX: 904136
41     INFORMATION FOR SEQ ID NO: 2:
42     SEQUENCE CHARACTERISTICS:
43     LENGTH: 503 amino acids
44     TYPE: amino acid
45     TOPOLOGY: linear
46     MOLECULE TYPE: protein
47     US-08-478-882-2

```

Query Match	3.8%	Score 89.5:	DB 2,	Length 503:	
Local Similarity	18.4%	Pred. NO.1.1:			
Matches 57:	Conservative	32:	Mismatches	96:	Indels 125; Gaps 14.
QY	40	ENDLQDNELPIDVOSASTDPANPLDEHPELYTALLENKTMILNCALNODIMRL	99		
Db	169	QEDIDASNIIDEDVSS-----GSTLEKSTPGYILHTDLPt---SQPTGRDR-----	212		
QY	100	ACYDTLVNGEPRAVLYKTKRST-----RLDETINOTIGKPOVYIQETTDPLFLMGNGK	153		
Db	213	---DAFFIGSTLATTATTPWVSANHTKONDERQWNPRIHSPVLLQDTT-----	258		
QY	154	MUTKDAKOLEVAAKQFPTPLSLFDLRNNPL-----WSSRHPNPMVYLPIFMHGKPNRS	209		
Db	259	-----RMt-----DIDRNSTAHGENMTQEQPP-----	283		
QY	210	PNTPSHEAKQFTPNFNRARBLKFQVSVYKAAEDLW-----GTDSDLMFGYTOOS	259		
Db	284	NNHHYQODEEERPH-----ASTTWADPNSTTEEAATQKEKMF-----EN	322		
QY	260	HMQITNGKN-----SRPRVADYQETILTQPVYSDLM-----DGKVMKIGMG	303		
Db	323	EMQ---GKNPRTPSDESHVTEGTSASAHNNHNSQKRWITQSOEDVSMTDFEDRISHPMGOG	379		
QY	304	AVVHNSNGESA	313		
Db	380	HOTESKGHSS	389		

```

1      RESULT 10
2      US-07-717-331F-2
3      ; Sequence 2, Application US/07717331F
4      ; Patent No. 5484905
5      ; GENERAL INFORMATION:
6      ; APPLICANT: June Nasrallah; Michael Nasrallah; and Joshua
7      ; APPLICANT: Stein
8      ; TITLE OF INVENTION: A Receptor Protein Kinase Gene
9      ; TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus
10     ; NUMBER OF SEQUENCES: 9
11     ; CORRESPONDENCE ADDRESS:
12     ; ADDRESSEE: Yahwak & Associates
13     ; STREET: 25 Skytop Drive
14     ; CITY: Trumbull
15     ; STATE: Connecticut
16     ; COUNTRY: USA
17     ; ZIP: 06611
18     ; COMPUTER READABLE FORM:
19     ; MEDIUM TYPE: Floppy Disk
20     ; COMPUTER: Macintosh
21     ; OPERATING SYSTEM: MS-DOS
22     ; SOFTWARE: Microsoft Word
23     ; CURRENT APPLICATION DATA:
24     ; APPLICATION NUMBER: US/07/717,331F
25     ; FILING DATE: June 19th 1991
26     ; CLASSIFICATION: 800
27     ; ATTORNEY/AGENT INFORMATION:
28     ; NAME: George M. Yahwak
29     ; REGISTRATION NUMBER: 26, 824
30     ; TELECOMMUNICATION INFORMATION:
31     ; TELEPHONE: (203)268-1951
32     ; TELEFAX: (203)268-1951
33     ; INFORMATION FOR SEQ ID NO: 2:
34     ; SEQUENCE CHARACTERISTICS:
35     ; LENGTH: 857 amino acids
36     ; TYPE: amino acid
37     ; TOPOLOGY: linear
38     ; MOLECULE TYPE: peptide
39     ;
40     ; US-07-717-331F-2

```

Query Match 3.88; score 89.5; DB 1; Length 857;

Best Local Similarity 21.1%; Pred. No. 2.7;
Matches 75; Conservative 51; Mismatches 110; Indels 119; Gaps 23;

```

QY 105 LVHGETPAVITKRSIRLDETTWOTIKGKPOVYQETTDPIFLMG---NEK---GMLTK 157
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 24 LIRPALSIYINT-----LSSTESLITISSNKTIV---SPGSIFEVGFPRTSRWLGWMYK 75
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 158 K-DAKOLEYAKQFTPLSLSF-----LDRNNTPLMSSRPINPVYLPPIFMHGK 205
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 76 KVSDFRYVWVANDPNPLSNIGTLKISGNMLVLLDHSNKKVWMTIN-----LTRGN 125
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 206 PNRSP-----NTPSHEAKQFT-----PNEFRAPELKQVSVKVAADLM 245
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 126 -ESSPVAAELLANGNFPMRSSNNNDASEYIMQSFDPDTDLPEMKIGYMLKT----- 177
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 246 GTDSDLMFGYTQOSHMOIFNGKNSRPFVHDYQ-----PEIFLQV---PVYSDLPMDGK 296
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 178 GLNRFL-----TSMRSSDPSSGNF---SYKLETSQSLPEFYLSRENPFMRHSGPWNG- 226
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 297 VRNIGMGAVH-----HSGCESAKLSRSNNRAV-----LMAGHEKMLTYMP--RI 339
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 227 IIRSGIPEDDQLSYMYNFIENNEEVAAYTFRMTNNSFYSLTLISEGYFORLTWYPSIRI 286
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 340 WGRIFEGSGSGOPD-----NPDLIDYGYGVDVRELYQLEKSNISCTVR-YNPRS 389
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 287 WNFNRPSPVDRQCDYIMGCP-----YAYCDV-----NTSPVCNCIQGFNPRN 329
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 11
US-08-633-879C-2
Sequence 2, Application US/08633879C
Patent No. 5928922

GENERAL INFORMATION:
APPLICANT: Kivirikko, Kari I.
APPLICANT: Pihlajaniemi, Taina
APPLICANT: Helekoski, Tarja I.
APPLICANT: Anonen, Pia P.
APPLICANT: Nissi, Ritva K.
APPLICANT: No. 5928922elainen, Minna K.
TITLE OF INVENTION: '2 SUBUNIT OF PROLYL-4-HYDROXYLASE
TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING SUCH SUBUNIT AND
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds, LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2811

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSEO for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/633,879C
FILING DATE: 10-APR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:
NAME: Adams, Samuel B
REGISTRATION NUMBER: 30,605
REFERENCE/DOCKET NUMBER: 8389-0041-999
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 537 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-633-879C-2

Query Match 3.7%; Score 87; DB 2; Length 537;
Best Local Similarity 21.0%; Pred. No. 2.3;
Matches 98; Conservative 53; Mismatches 146; Indels 170; Gaps 25;

```

QY 1 MKVSLSTLTLSISCFALIAQQAQAVNPVAFVDESRENDGQDMLPDVQSA----- 56
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MKLQVLVLVL-LMSWFGVLSMVOAEFTTSIGHMTDLIYAEQDLVQSIKEYILVEAKLAK 59
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 57 -----TOSASVD-----TANPLDEH-----PELYTALENKT--MLNCS 90
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 60 IKSWSKMEALTSRSADPEGYLAHPVNAVYKLVKRLMTDMPALGDVLYQDASAGFVANLS 119
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 91 ALNQ-----DIMRLA-CY---DTLVHGETPAVITKRSIRLDETI----- 126
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 120 VQROFFPTDESGARAIARLADPTVYKLDPTLSRGELPCT-KYQAMLSYDDCGGLGRSA 178
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 127 -----WOTIKGKQVYQETTDPIFLMGNEKMLTKK-DAKOLEYAKQFTPLSLSF 179
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 179 YNEGDIYHTVLMMEQVLIKQ-----LDAGEEATVTSVLVDYLSYVFGQGLHRAVEL 231
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 180 DRNNTPLMSSRPINPVYLPFIHMGKPNRSPNPSHEAK-----OFTNPFAPAPLKF 232
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 232 TRRLSL-----DPSHERAGNLRKYFERLLEEGKSLSN 266
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 233 QVSVKYKAAEDLMGTDSDLMFGYTQOSHMOIFNGKNSRPFVHDYQEIFLQPVYSDLP 292
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 267 QTDAGLATQENLY-----ERP---FDYLPD-----RDVYESL- 295
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 293 WDKVVRNIGMGAVHHSNGESAKLS-RSMNRAVYLM--AGMEKMLTYMPRIHGRIFKSGS 349
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 296 -----CRGEGKILPBRQKLFGRYHNGNRVPOLLIP-----FKR-- 331
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 350 SOPDDNPDLIDYGYGVDVRELYQLEKSNIS-----GTVRNPYRSG 390
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 332 EDEWDSPIYVRY---DYMSDEIERIKETAKPKLAATYR-DKTKG 374
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 12
PCT-US96-03916-6
Sequence 6, Application PC/TUS9603916
GENERAL INFORMATION:

APPLICANT: Wild, Martha A.
APPLICANT: Cochran, Mark D.
TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/03916
FILING DATE: 23-MAR-1995
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,597
FILING DATE: 24-SEP-1993

```

      ATTORNEY/AGENT INFORMATION:
        NAME: White, John P.
        REGISTRATION NUMBER: 28,678
        REFERENCE/DOCKET NUMBER: 39116-A
        TELECOMMUNICATION INFORMATION:
          TELEPHONE: (212) 278-0400
          TELEFAX: (212) 391-0525
        INFORMATION FOR SEQ ID NO: 6:
          SEQUENCE CHARACTERISTICS:
            LENGTH: 985 amino acids
            TYPE: amino acid
            TOPOLOGY: linear
          MOLECULE TYPE: protein
PCT-US96-03916-6

Query Match              3 7%; Score 87; DB 5; Length 985:
Best Local Similarity 19.4%, Pred. No. 6.1; Mismatches 126; Indels 130; Gaps 20
Matches 73; Conservative 47;

QY    48 ELPIVQASATQSASTDTANPLDHEPELYT-TALEKTKMLINCSALNODIMRLACYDTLV 106
Db    563 QTPSVPRPTLTSSR-----BPALTFRQSGAGTEAFQTGSAEPRTMTQSTETHF 613
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY    107 HGETPAVIKTAKSISILDETITWOTIGKPOVYIQE--TTDI---FLMGNKGMITKKDAKO 162
Db    614 FTQAISTVPKAP-----QTSPSTEPLVLSQSPSTEEVPETRI LGAEIETQTGPSA-- 662
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY    163 LEVAKQFTPLSLSPDLDDNNTPLMWSRR-----HN-----PMYVLP----IPMHCK 205
Db    663 ---AEVYTRRSSTMPETAQSTPLASQNSTSGTGHTNTEPRPYQTTPHOKLITYENK 719
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY    206 PKRSNTPS--HE---AKQFP--NEFRADKLFOYSVKYA-----AEDIWGTD 248
Db    720 TLFSPTVVSEFEHESMTAESQTPLLVKIVYEKFNSDGEYATACVSIVKYRETMKVVD 779
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY    249 -----SDLDFGYTOOSHQIFNGKNS-----RPF----- 272
Db    760 LYDVWDELTSGNPSAGVFNSNEKW---QKOLYRVTDGRISVOLMCLSCSHSPEPTCLFD 836
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY    273 ----RVHDYOEIFLT--QPVSDLPWDGV-----RMIGKAVAHNSN 309
Db    837 TSLIAREKDIADELYTSDPQTAYCITLPSGVPRFEWSLNNVSLPEYLTLATTVVSHTA 896
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY    310 GE-----SAKLSRW 319
Db    897 GGSTVWKSSARAGEAM 912
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
PCT-US96-03916-66
Sequence 66, Application PC/TUS9603916
GENERAL INFORMATION:
APPLICANT: wild, Martha A.
APPLICANT: Cochran, Mark D.
TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSER: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/03916
FILING DATE: 23-MAR-1995

```

```

: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/126,597
: FILING DATE: 24-SEP-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: White, John P.
: REGISTRATION NUMBER: 28,678
: REFERENCE/DOCKET NUMBER: 3116-A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 278-0400
: TELEFAX: (212) 391-0525
: INFORMATION FOR SEQ ID NO: 66:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 985 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: PCT-US96-03916-66

:
:
: Query Match 3.7%, Score 87; DB 5; Length 985;
: Best Local Similarity 19.4%; Pred. No. 6.1;
: Matches 73; Conservative 47; Mismatches 126; Indels 130; Gaps 20.

:
: QY 48 ELIPDVQATQSASDTDTANPLDENHEPELTY-TALENKTMLINCSALINDIMRLACTDLY 106
: Db 563 QTPSTVPPTTLSSST-----EPAIFTRTQSAGTEAFQTSSAPDPMRTGSTETHF 613
: QY 107 HGEPAVAVTKTRSRIRLDEITIMOTIKRGKVOIVQE--TTDPIT--FLMGNEKGMITKKDAKO 162
: Db 614 FTQASSTVPKAT-----QTPTSEPEVLQSPSTSEPEVPTTRLGAPPELTQTPRSA-- 662
: QY 163 LEVAAKQFTPLSLSDLDLRNNTPLMSSRP-----HN-----PMYVLP----IFMHCK 205
: Db 663 --AEVYTRSSSTMPETAQSTPLASOMPTSSGCTHNTPEPTYPVQTPPHQKLYTENK 719
: QY 206 PNRSPNTS--HE--AAQGFQ--NEFRAPELKEQYSVKYKA-----AEDLMGND 248
: Db 720 TLSPTTVVSEHEHMTAESQTPILDVKIYEVKFSNDGEVTAFCVSTVSPRYEVLNWKVD 779
: QY 249 -----SDLMFGYTOOSHMOIFNGKNS-----RPF---- 272
: Db 780 LVDVMDDEISGNPAGVFNSEKWK--QKQLYRVYDGRTRSYQLMGLSCTSHSPEPECLFD 836
: QY 273 -----RVNDYQPEILTY--QPVYSDLPWDKAY-----RMIGMGAVHNHN 309
: Db 837 TSLIAREKDIAELFTSDOPQAYCTITLPSGVVPREFEWSLNNVSLPEYLATATVYVSHR 896
: QY 310 GE-----SAKLSRSM 319
: Db 897 GQSTVWKSSARAGEAW 912

:
: RESULT 14
: US-08-961-083-2
: Sequence 2, Application US/08961083
: Patent No. 6159469
: GENERAL INFORMATION:
: APPLICANT: Choi et. al.
: TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
: NUMBER OF SEQUENCES: 452
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
: COMPUTER: HP Vectra 486/33
: OPERATING SYSTEM: MSDOS version 6.2
: SOFTWARE: ASCII Text

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:18:32 ; Search time 44.45 Seconds
757.461 Million cell updates/sec

Title: US-09-787-083-4

Perfect score: 2360

Sequence: 1 MKVSLSTLTLSTLPCFAILA.....YNHEATSFVGLKNDMMGL 442

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	815	34.5	382	2	EB1195
2	815	34.5	409	2	HB1831
3	373	15.8	286	2	B36971
4	366	15.5	329	2	D81279
5	358	15.2	289	2	A36971
6	357	15.1	289	1	PSECA1
7	357	15.1	289	2	EB6069
8	321.5	13.6	289	2	C36971
9	246.5	10.4	355	2	C64582
10	240.5	10.2	355	2	H71930
11	109.5	4.6	824	3	JC7532
12	109	4.6	800	2	A29003
13	109	4.6	822	2	JT0611
14	106.5	4.5	602	1	TVTRR
15	101.5	4.3	783	2	JC5467
16	98.5	4.2	1658	2	T42642
17	97.5	4.1	797	2	T46737
18	95.5	4.0	389	2	S76490
19	94.5	4.0	397	2	A35136
20	94.5	4.0	601	2	T26062
21	94.5	4.0	655	2	T26061
22	94.5	4.0	1641	2	D82704
23	94	4.0	696	2	S55694
24	94	4.0	765	2	T35719
25	93.5	4.0	564	2	T40777
26	93.5	4.0	901	2	T20122
27	93	3.9	633	2	C81956
28	93	3.9	791	2	H96839
29	93	3.9	4273	2	C69679

30	93	3.9	4307	2	T20721
31	92.5	3.9	857	1	A41369
32	92.5	3.9	1807	2	T30940
33	92	3.9	324	2	B69521
34	92	3.9	888	2	S50801
35	92	3.9	2971	2	T08026
36	92	3.9	5005	2	F82884
37	91.5	3.9	796	2	JC7355
38	91	3.9	367	2	T24058
39	91	3.9	1078	2	T18352
40	91	3.9	1474	1	MAHU
41	90.5	3.8	386	2	A81328
42	90.5	3.8	467	1	A49377
43	90	3.8	841	2	JC5894
44	90	3.8	1788	2	T29043
45	89.5	3.8	477	2	S23257

ALIGNMENTS

RESULT 1	
EB1195	phospholipase A1, probable phospholip
C:Species: Neisseria meningitidis	
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001	
C:Accession: EB1195	
R:Retelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B. ri, H.; Qin, H.; Yamathavan, J.; Gill, J.; Scariato, V.; Masigian, V.; Pizzza, M. Science 287, 1809-1815, 2000	
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappaport, R.; A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58. A:Reference number: A81000; MUID:20175755	
A:Accession: EB1195	
A:Status: preliminary	
A:Molecule type: DNA	
A:Residues: 1-382 <TE>	
A:Cross-references: GB:AE002403; GB:AE002098; NID:g7225688; PIDN:AAFA0901.1; PID:g7222	
A:Experimental source: serogroup B, strain MC58	
C:Genetics:	
A:Gene: NMB0464	
Query Match	34.5%; Score 815; DB 2; Length 382;
Best Local Similarity	44.7%; Pred. No. 1.4e-59;
Matches 163; Conservative 67; Mismatches 111; Indels 24; Gaps 8;	
OY 87 INCSALNODIMRLACVDTLVHGETPAVI-----KTKRSIRLDEITMOT-KGKPOVYQE 140	
DB 33 LOCALITDNTVTLACYDRFAAQLPSAGQEQEQSKAVNLTEYRSSLDKEAVIYVEK 92	
OY 141 TTDPIFLMGNEKGMLETKDAQLEYAAQFTPLSLFSDLRNN-TPLMSSRPHNPYVLP 199	
DB 93 GGDAL-----PADSAGETADITYTPLSLMYDLDRKLDKGLGVREHNPYVLP 139	
OY 200 IYHMGKPNRSPYTPSHEAR-QFTPNDFRAPLEKFOYSAKYKAAEDLMGDSULMTCYQQ 258	
DB 140 IMYNNSPYAPSPGSPRTGVQOKRAETKLYVSFSKJAEDLFKTRALDMLGCTYOR 199	
OY 259 SHMOJFN-GKNSRPFVRVHDOPEIFLTQPVYSDLPMDGVKVRIGAGVAHNSGESAKTSS 317	
DB 200 SMQIYNOGRKAPPRANDYKPEITLTQPVKADLPFGGLRMAGFPVHQSOGSRPESR 259	
OY 318 SWNRAYLMAEMKMLVWPRITWGRIFKESGSGQDPDNDIIDDYGYGQVRLYOLENKS 377	
DB 260 SMNRITAYMAAGMEKGLTVIPRVWVRAFDQ-SSDK-NDNPDIDADYGYGQVRLYRLNDQ 317	
OY 378 NISGIVRINPRSGKALQLDIYVPLKGISGFGYFQYGGSLIDYNEHATSFVGLMIN 437	
DB 318 NVYSVLRYNPKTYGAIETAAVTFPIKGLKGVVRGFHGESGLIDYNNKONGIGILMFN 377	
OY 438 DMNGL 442	

QY 313 -----AKSRSMNRAVLAGMEKNTLTPRIMGR 342
Db 424 VEDVVIENAGALKLGLSDASNDVSEGNWYANRLSADGWCK 465

RESULT 13

JT0611
cellulase (EC 3.2.1.4), alkaline - Bacillus sp. (strain KSM-64)
N:Alternate names: endo-1,4-beta-glucanase
C:Species: Bacillus sp.
C>Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 10-Dec-1999
C:Accession: JT0611
R:Sumitomo, N.; Ozaki, K.; Kawai, S.; Ito, S.
Biosci. Biotechnol. Biochem. 56, 872-877, 1992
A:Title: Nucleotide sequence of the gene for an alkaline endoglucanase from an alkalophilic
A:Reference number: JT0611; MUID:92305459
A:Accession: JT0611
A:Molecule type: DNA
A:Residues: 1-822 <SUM>
A:Cross-references: GB:M84963; NID:9289264; PIDN:AAA73189.1; PID:9289266

C:Function:
A:Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as ce
A:Pathway: cellulose degradation
C:Superfamily: Bacillus sp. alkaline cellulase; Thermotoga xylinase A amino-terminal re
C:Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation
F:585-726/Domain: Thermotoga xylinase A amino-terminal repeat homology <TXA>

Query Match 4.6%; Score 109; DB 2; Length 822;
Best Local Similarity 19.7%; Pred. No. 1;

Matches 91; Conservative 51; Mismatches 150; Indels 170; Gaps 24;

QY 30 PVAFVDEVRSKND-----IGQDNELLIGVOSATQASDTDPANPL-DEHEPELY-----TT 78
Db 25 PTLAEGNTRNFHILGNDNVKRPSEAGALQLEVDQMTLVQHGKIKLRGSMTH 84
QY 79 ALENKMYLINCAL-----NODIMRLACY-DLVHGETPRAVIKTKRSIRLDETT----- 126
Db 85 GLQWPEELINDNAYKALANDWESNMIRLAMYVGENGYASNPBLIKSVIKIGDILAIENDM 144
QY 127 -----MGT-----IKGKPOVYVOETTP-----IFLMGN 150
Db 145 YVAVDMHVNARCPDRPVYAGAEDFFRDIALYPPNPHITIELANEPSSNNNGAGAPNN 204
QY 151 EKGMLTRKDAKQLEYAKOFTPLSLSFDDRN-----NTPLMSSRP-----HNPMY 196
Db 205 EEEGMNWK-----EYADPIYEMLRDSGNADNIIIVGSPWMSQRPDLAADNPIDDHHTWY 259
QY 197 VLPIF--MGKPNRS--PNTPSHEARQFTPNFRAPELKFQVSVKKAEDLNGT----- 247
Db 260 TVHFYGSNHAASYEPPEPNSERGVNMSNTRYA-----LENGVAVFATE--WGTSQANG 313
QY 248 -----DSDLMFGYQQ-----SHMOIFNGKNS-----RPFVYH----- 275
Db 314 DGGPIYDEADVWLEFLNENNISWANSLTN-KNEVSGAPTFEELGKSNATSLDPEPDQVW 372
QY 276 -----DYQPEIFLTQPVYSDLPW---DGKVRMIGMGAVHSHNGES- 312
Db 373 VPPELSLSGEYVARIKGVYEP---IDRTKYTKVLMDFNDGTRKQGGV-----NDDSP 423
QY 313 -----AKSRSMNRAVLAGMEKNTLTPRIMGR 342
Db 424 VEDVVIENAGALKLGLSDASNDVSEGNWYANRLSADGWCK 465

RESULT 14

TVKTRR
protein kinase (EC 2.7.1.37) raf - rat
N:Alternate names: kinase-related transforming protein raf; raf proto-oncogene protein-s
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 11-Jun-1999
C:Accession: B26126

R: Ishikawa, F.; Takaku, F.; Nagao, M.; Sugimura, T.
Mol. Cell. Biol. 7, 1226-1232, 1987
A:Title: Rat c-raf oncogene activation by a rearrangement that produces a fused prote
A:Reference number: A26126; MUID:87172791
A:Accession: B26126
A:Molecule type: mRNA
A:Residues: 1-602 <TSH>
A:Cross-references: GB:M15428; NID:9206546; PIDN:AAA42002.1; PID:9206547

C:Genetics:
A:Gene: raf
C:Superfamily: rat protein kinase raf; protein kinase homology
C:Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; serine/thre
F:501-567/Domain: protein kinase homology <KIN>
F:509-317/Region: protein kinase ATP-binding motif
F:329/Active site: Lys #status predicted
F:453/Binding site: phosphate (Ser) (covalent) (by autophosphorylation) #status predi

Query Match 4.5%; Score 106.5; DB 1; Length 602;
Best Local Similarity 19.0%; Pred. No. 1;
Matches 92; Conservative 74; Mismatches 172; Indels 145; Gaps 24;

QY 23 QAQAVPVAFVDEVRSKNDLIGQDNELLIGVOSATQASDTDPANPLDEHEPELYTTALEN 82
Db 81 ELEKLNQVAVLEKKKELETAQDRNL--GIQGFTRAKE-----LEAKRDLIRT---N 131
QY 83 KTMILNCASALNODIMRLACYDTLVHGETPRAVIKTKRSIRLDETTWO--TIKGRPOVYOE 140
Db 132 ERLSQEVEYLTEDEVKRL--NEKLKESNT--TKGEIQLKIDELQASDVVYKREKRLQOE 186
QY 141 -----TTPPIFLMGNEKG-----MLTKKDAKQLEYAKOFTPLS 174
Db 187 KELLHQNQSNLNLTKTKTDELALGREGKNEILELTKLENKEEDAIRSHSESASPSA 246
QY 175 LSFDDLRNNTPLMSSRPNHNVYLPIFMHGKPNRSPNTPSHEARQFTPNFRAPELKFQV 234
Db 247 LS--SSPNLSPFGSQPKTP-----VPAQREARASGQENKRIKRPQGRSSVYWEI 298
QY 235 SVKVAKAEDMGTDSDLMFGYQQSHW-----QIFNGKNSRPFVYHVOPEI----- 281
Db 299 EASEVMLSTRIGSGS--FGTVYKKGWHDVAVKILKVDPRTPEQLQAFNEVAVLKTR 355
QY 282 -----FLTQ-----PVYSDL--PMGKVMIT-----GMAVNH 307
Db 356 HVNILLFMGYTRKDNLAIVQWCEGSSLYKHLHVQETKFMQFLDIARQTAGOMDYLAH 415
QY 308 SNGESAKLSRSMNRAVLAGMEK-----NLTPRIMGRIFKESGSG----- 351
Db 416 KNIIHRDKMS--NNIFLHEGLTVKIGDFGLATYKSR-----SSQVYEQPTGSVLYM 466
QY 352 -----PDDNPDL--DYGYGDVREFLYQLENKSNISGTVRVNPBSGKALQLDVYVPL 402
Db 467 APEYIRQDNNPFSFGSDVSYGIV--LYEL-----MTGELPYSHINNROI-----IFWV 515
QY 403 GKG 405
Db 516 GRG 518

RESULT 15

JCS467
cellulase (EC 3.2.1.4) - Bacillus sp.
N:Alternate names: Endo-1,4-beta-glucanase
C:Species: Bacillus sp.
C>Date: 07-Jul-1997 #sequence_revision 18-Jul-1997 #text_change 10-Dec-1999
C:Accession: JCS467
R: Miyatake, M.; Imada, K.
Biosci. Biotechnol. Biochem. 61, 362-364, 1997
A:Title: A gene encoding endo-1,4-beta-glucanase from Bacillus sp. 22-28.
A:Reference number: JCS467; MUID:97212038
A:Accession: JCS467
A:Molecule type: DNA
A:Residues: 1-783 <MTY>

A:Cross-references: DDBJ:D85236

A:Experimental source: strain 22-28

A:Note: neither the complete nucleic acid sequence nor the complete translation are shown
C:Superfamily: Bacillus sp. alkaline cellulase; Thermotoga xylosanase A amino-terminal ref

C:Keywords: glycosidase; hydrolase
F:587-728/Domain: Thermotoga xylosanase A amino-terminal repeat homology <TXA>

```
Query Match          4.3%; Score 101.5; DB 2; Length 783;
Best Local Similarity 19.7%; Pred. No. 3.9;
Matches 91; Conservative 54; Mismatches 148; Indels 169; Gaps 26;

QY 30 PYAFVDEYRSKND-----LGDNELLIGVQATQSSTDYANPL-DEHEPELY-----TT 78
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 25 PTALAEENGTREDNDPDLGNGENVRKRPSEAGALQKEVDGQMTLVDOHGEEKIQLRGMSTH 84
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 79 ALENKTMILNCSAL-----NODIMRLACYDTLVHGE-----TPAVIKTK-----RSI 120
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 85 GLQWPEELINDNAYKALSNDDMSNMIRLAMY-----VGNGVATNPBELIKQRYIDGIELAI 140
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 RLDETI---WQT-----IKGPOVYVQETTD-----IF 146
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 141 ENDMTVIYDWHVHAPGDPDPYAGAEDEFRDIALYPNPHIYELANEPSSNNNGAG 200
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 147 IMGNEKGMILTKKDAKQLEYAAKQFTPLSLSPDLRN-----NTPLMSSRP-----H 192
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 201 IPNNEEGMKAVK-----EXADPIVEMLRDGNADNIIIVGSPNMSQRPDLADNPINDH 255
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 193 NPMVYLPIF--MHGKPNRS--PNTPSHEARQFTPNFRAPELKFQVSVKKAEDLMGT- 247
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 256 HTMYTVHFTYTGSHASTESYPPEPTNSEKGNVMSNTRYA-----LENGAVVFATE--WGTIS 309
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 248 -----DSDLMFGYTQO-----SHWQIFNGKNS-----RPRV----- 274
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 310 QANGDGCPYDFDEADVIEFLNENINISMANWSLTN--KNEVSGAFPPFELGKSNATSLDGP 368
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 275 -----HDYQPELFLQPYVSDLPW---DGKVRMIGMGA----- 304
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 369 DQVMAPEELSLSGEYVRAIRIKGAKYEP--IDRTRYTKVLMDFNDGTGKGFVNSDSPNK 425
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 305 ----VHHSNGESAKLSRSMNRAYLMAGMEKKNITVMPRIWGR 342
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 426 EAIEYVENENG-TLRISGLNVSNLSDGNFMANFRLSANGMCK 466
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: November 30, 2001, 14:18:34
Job time: 256 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:27:00 ; Search time 28.02 Seconds

(without alignments)
578.367 Million cell updates/sec

Title: US-09-787-083-4
Perfect score: 2360
Sequence: 1 MKVSLSTLTLTLPCFALIA.....YNHEATSFVGLMLDMNGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 segs, 36664827 residues
Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	373	15.8	286	1	P37446 KLEPN
2	358	15.2	289	1	P37446 KLEPN
3	357	15.1	289	1	P37446 KLEPN
4	321.5	13.6	289	1	P37446 KLEPN
5	109	4.6	800	1	P37446 KLEPN
6	94.5	4.0	397	1	P37446 KLEPN
7	93	3.9	4273	1	P37446 KLEPN
8	92.5	3.9	849	1	P37446 KLEPN
9	92	3.9	324	1	P37446 KLEPN
10	92	3.9	696	1	P37446 KLEPN
11	92	3.9	888	1	P37446 KLEPN
12	91.5	3.9	969	1	P37446 KLEPN
13	91	3.9	496	1	P37446 KLEPN
14	91	3.9	1474	1	P37446 KLEPN
15	90.5	3.8	467	1	P37446 KLEPN
16	90	3.8	1158	1	P37446 KLEPN
17	90	3.8	1788	1	P37446 KLEPN
18	89.5	3.8	486	1	P37446 KLEPN
19	89.5	3.8	669	1	P37446 KLEPN
20	89.5	3.8	1478	1	P37446 KLEPN
21	89.5	3.8	1840	1	P37446 KLEPN
22	89	3.8	1426	1	P37446 KLEPN
23	88.5	3.8	825	1	P37446 KLEPN
24	88.5	3.8	992	1	P37446 KLEPN
25	88	3.7	810	1	P37446 KLEPN
26	88	3.7	1115	1	P37446 KLEPN
27	88	3.7	1382	1	P37446 KLEPN
28	87.5	3.7	844	1	P37446 KLEPN
29	87.5	3.7	1379	1	P37446 KLEPN
30	87	3.7	503	1	P37446 KLEPN
31	87	3.7	666	1	P37446 KLEPN
32	87	3.7	842	1	P37446 KLEPN
33	87	3.7	882	1	P37446 KLEPN

ALIGNMENTS

34	86	3.6	882	1	HSS1_HUMAN	P52848 homo sapien
35	86	3.6	1087	1	XYNX_CLOTM	P38535 clostridium
36	85.5	3.6	353	1	DCUP_BACSV	P23395 bacillus su
37	85.5	3.6	790	1	SEIL_MOUSE	Q92426 mus musculu
38	85.5	3.6	794	1	SEIL_HUMAN	Q9ubv2 homo sapien
39	85	3.6	403	1	P37_MYCHR	P15363 mycoplasma
40	85	3.6	491	1	TYXH_PHASP	P11982 phasianidae
41	85	3.6	560	1	DTXH_CORBE	P00589 corynebacte
42	85	3.6	828	1	BGAL_BRAOL	P49676 brassica ol
43	85	3.6	1184	1	ALAC_ARATH	P57792 arabidopsis
44	85	3.6	2231	1	SENI_YEAST	Q00416 saccharomyc
45	84.5	3.6	478	1	DHGB_ACICA	P13650 actinobact

RESULT 1
ID PAL_KLEPN STANDARD: PRT: 286 AA.
AC P37446;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT
DE PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-
DE ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).
GN PLDA.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_Taxid=573;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94131966; PubMed=8300539;
RA Brock R.G.P.M., Brinkman E., van Boxtel R., Bekkers A.C.A.P.,
RA Verheij H.M., Tomassen J.;
RT "Molecular characterization of enterobacterial plda genes encoding
RT outer membrane phospholipase A.";
RL J. Bacteriol. 176:861-870(1994).
CC -1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE
CC A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.
CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES
CC LOCATED THERE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X76901; CA54223.1; -
CC PIR: B36971; B36971.
CC PIR: S40129; S40129.
CC InterPro: IPR003187; PLA1.
CC Pfam: Pf02253; PLA1; 1.
CC Hydrolyase; Lipid degradation; Outer membrane; Signal; Calcium.
CC SIGNAL 20
CC CHAIN 21 286 BY SIMILARITY.
CC ACT_SITE 161 161 PHOSPHOLIPASE A1.
CC SEQUENCE 286 AA; 3E39F863085108A3 CRC64;

Query Match 15.8%; Score 373; DB 1; Length 286;
Best Local Similarity 42.1%; Pred. No. 3.3e-23;
Matches 90; Conservative 27; Mismatches 87; Indels 10; Gaps 5;

Query Match 15.2%, Score 358, DB 1, Length 289;

Best Local Similarity 39.3%; Pred. No. 5.6e-22; Matches 84; Conservative 30; Mismatches 90; Indels 10; Gaps 4;

```
OY      226 RAPELKFQVSVKVKRAEDIMGCTDSLMFGCYTQSIMQIFNGKNRPFRRHDTQPFIETQ 285
       + : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      83 RDEDFKFLQLSLFPLMRIGLGSPNSVLGASYSYOKSMWOLSNSESSPFRFTNEPOLFLGF 142
OY      286 PV-YEYDLPWDGCVRIRIGAGVAHHNNGESAKTLSRSWNRAVILAMGMKNTLVPRITMGRIE 344
       + : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      143 ATDIYFAQM--TLRVENKGYNHDSNGRDPSRTYNNRLTYRLMAENGMLVEYKRMYVI- 199
OY      345 KEGSGSQPDNDNDLLDYGYGDVRFLEYOLENKSNISGTIVRYNPRSGKALQLDIYVPGLCK 404
       ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      200 -----GSTDDNMEDIIRKYMGYYOLKIGYHL-GCAVYSANKQYMMNTGCGAEEGLSYPVK 253
OY      405 GISGGFOJIEOGYGOSLIDYNHEATSFVGVMIND 438
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      254 HVLITYOVSGTGESLIDYNFQTRVGVGVLEMD 287

RESULT      3
PAL_ECOLI ID PAL_ECOLI STANDARD: PRT: 289 AA.
AC P00631. 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DERGENT-RESISTANT
DE ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).
GN PLDA OR B3821 OR Z5342 OR ECS4751.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CX NCBI_Taxid=562, 83334;
RX [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=85157492; PubMed=6397464;
RX Homma H., Kobayashi T., Chiba N., Karasawa K., Mizushima H., Kudo I.,
RA Inoue K., Ikeda H., Sekiguchi M., Nojima S.;
RA "The DNA sequence encoding plda gene, the structural gene for
RL detergent-resistant phospholipase A of E. coli.";
RT J. Biochem. 96:1655-1664(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=92358234; PubMed=1379743;
RA Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;
RT "Analysis of the Escherichia coli genome: DNA sequence of the region
RL from 84.5 to 86.5 minutes.";
RN Science 257:771-778(1992).
RN [-3]
RP REVISION TO 14-15.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregory J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RN Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., DiMantana E.T., Potamouis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
```

RL Nature 409:529-533(2001).
 RN (5)
 RP SEQUENCE FROM N.A.
 RC STRAIN-0157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tohe T.,
 Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 Kuhara S., Shiba T., Hattori M., Shingawa H.;
 RT Complete genome sequence of enterohemorrhagic *Escherichia coli*
 O157:H7 and genomic comparison with a laboratory strain K-12.*;
 RL DNA Res. 8:11-22(2001).
 RN (6)
 RP SEQUENCE OF 30-289 FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=85003590; PubMed=6383820;
 RA de Gaus P., Verheij H.M., Riegman N.H., Hoekstra W.P.M., de Haas G.H.;
 RT "The pro- and mature forms of the *E. coli* K-12 outer membrane
 phospholipase A are identical.";
 RL EMBO J. 3:1799-1802(1984).
 RN (7)
 RP SEQUENCE OF 174-289 FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=87115164; PubMed=3027506;
 RA Irimo N., Nakayama K., Nakayama H.;
 RT "The *recG* gene of *Escherichia coli* K12: primary structure and
 evidence for SOS regulation.";
 RL Mol. Gen. Genet. 205:298-304(1986).
 RN (8)
 RP MUTAGENESIS OF SER-172.
 RX MEDLINE=94131966; PubMed=8300539;
 RA Brok R.G.P.M., Brinkman E., van Bortel R., Bekkers A.C.A.P.,
 Verheij H.M., Tommassen J.;
 RT Molecular characterization of enterobacterial *plda* genes encoding
 outer membrane phospholipase A.*;
 RL J. Bacteriol. 176:861-870(1994).
 RN (9)
 RP ACTIVE SITE SER-164, AND PARTIAL SEQUENCE.
 RX MEDLINE=91249806; PubMed=2040286;
 RA Horrevorts A.U.G., Verheij H.M., de Haas G.H.;
 RT "Inactivation of *Escherichia coli* outer-membrane phospholipase A by
 the affinity label hexadecanesulfonyl fluoride. Evidence for an
 active-site serine.";
 RL Eur. J. Biochem. 198:247-253(1991).
 CC -1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE
 A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.
 CC REQUIRED FOR EFFICIENT SECRETION OF BACTERIOCINS, SEEMS TO BE
 CC DORMANT IN NORMAL GROWING CELLS.
 CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-
 CC PHOSPHOCHOLINE + A FATTY ACID ANION.
 CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-
 CC PHOSPHOCHOLINE + A FATTY ACID ANION.
 CC -1- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES
 CC LOCATED THERE.
 CC -1- INDUCTION: BY MEMBRANE DAMAGING, FOR EXAMPLE, BY PHAGE-INDUCED
 CC LYSIS OR TEMPERATURE SHOCK.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X02143; CAA26081.1; -;
 DR EMBL: M87049; AAA67617.1; -;
 DR EMBL: AE000458; AAC76824.1; -;
 DR EMBL: AE005613; AAG59017.1; -;
 DR EMBL: AP002567; BAB38174.1; -;
 DR EMBL: M30198; AAA24516.1; -;
 DR PIR: A00771; PSECA.

DR PIR: A22133; PSECA1.
 DR PIR: S30711; S30711.
 DR EcoGene; EG10738; *plda*.
 DR InterPro; IPR003187; *PLA1*.
 DR Pfam; PF02253; *PLA1*; 1.
 KW Hydrolase; Lipid degradation; Outer membrane; Signal; Calcium;
 KW Complete proteome.
 FT STGNAL 1 20
 FT CHAIN 21 289 PHOSPHOLIPASE A1.
 FT ACT_SITE 164 164
 FT MOTAGEN 172 172 S->F: INACTIVE PROTEIN.
 FT CONFLICT 14 15 LP -> FA (IN REF. 2).
 FT CONFLICT 30 33 DAPA -> MTRQ (IN REF. 6).
 SQ SEQUENCE 289 AA; 33163 MM; A688AD32AA60F218 CRC64;
 Query Match 15.1%; Score 357; DB 1; Length 289;
 Best Local Similarity 39.3%; Pred. No. 6,7e-22;
 Matches 84; Conservative 29; Mismatches 91; Indels 10; Gaps 4;
 QY 226 RAPELKROVSKYKAAEDLMGTSDLMFGYTOQSHWQIFGNKNSRPVRVHYDEIFLQ 285
 Db 83 RKDEVKQQLSLAPFLMRGLGPNVSLGASVYQKSWQLSNSESPPRETYEPQLFGLF 142
 QY 286 PV-VSDLPWDGKVRMIGGAVHNSNGESAKLSRSNNRAYLMAAGMKNKLTVMPTIRGRIF 344
 Db 143 ATDYRFAGW--TLRDVENGVNHSNDSNPTSRSMNRLYTLMAENGWLVYKPPWYV- 199
 QY 345 KEGSQPDNDNDILDYVYGVDFEFLYLENKSNSIGTVRYNPPRSKGALQLDVYPLGK 404
 Db 200 -----GNTDDKDPDITKTMGYQLKIGYHL-GDAVLSAKGQYMNMTGCGALGLSYPTTK 253
 QY 405 GISGYFOIFGQSLIDYNEHATSFGVGLMND 438
 Db 254 HVRLYTVYSGYSGESLIDYNFNQTRVGVMND 287
 RESULT 4
 PAL_PROVU STANDARD; PRT; 289 AA.
 AC P37447;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE PHOSPHOLIPASE A1 PRECURSOR (EC 3.1.1.32) (DETERGENT-RESISTANT
 DE PHOSPHOLIPASE A) (DR-PHOSPHOLIPASE A) (PHOSPHATIDYLCHOLINE 1-
 DE ACYLHYDROLASE) (OUTER MEMBRANE PHOSPHOLIPASE A) (OM PLA).
 GN *plda*.
 OS Proteus vulgaris.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Proteus.
 OC NCBI_TaxID=585;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94131966; PubMed=8300539;
 RA Brok R.G.P.M., Brinkman E., van Bortel R., Bekkers A.C.A.P.,
 Verheij H.M., Tommassen J.;
 RT Molecular characterization of enterobacterial *plda* genes encoding
 RT outer membrane phospholipase A.*;
 RL J. Bacteriol. 176:861-870(1994).
 CC -1- FUNCTION: HYDROLYSIS OF PHOSPHATIDYLCHOLINE WITH PHOSPHOLIPASE
 CC A2 (EC 3.1.1.4) AND PHOSPHOLIPASE A1 (EC 3.1.1.32) ACTIVITIES.
 CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERO-
 CC PHOSPHOCHOLINE + A FATTY ACID ANION.
 CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 2-ACYLGLYCERO-
 CC PHOSPHOCHOLINE + A FATTY ACID ANION.
 CC -1- COFACTOR: REQUIRES CALCIUM IONS FOR ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE; ONE OF THE VERY FEW ENZYMES
 CC LOCATED THERE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

```
CC      use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sdb.ch/announce). CC      or send an email to license@isb-sdb.ch).
CC      -----
DR       EMBL; X76902; CAA54224.1; -.
DR       PIR; S46971; C36971.
DR       PIR; S40130; S40130.
DR       InterPro; IPR003187; PLAL.
DR       Pfam; PF02253; PLAL; 1.
KW       Hydrolyase; Lipid degradation; Outer membrane; Signal; Calcium.
FT       SIGNAL                     1          20
FT       CHAIN                      21         289   BY SIMILARITY.
FT       ACT_SITE                   164        164   PHOSPHOLIPASE A1.
SQ       SEQUENCE   289 AA;  32944 MW;  D7516CFBBA06997 CRC64;

Query Match                               13.6%; Score 321.5; DB 1; Length 289;
Best Local Similarity 33.3%; Pred. NO.5.3e-19;
Matches 87; Conservative 42; Mismatches 115; Indels 17; Gaps 7;

QY    179 LDRNNTPLMSSRRHPNVPVLPIFMHGKFNRSRSEPTSHBAROFTPNRFAPELKFQVSKV 238
     + : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db    43 LOEHNP-FITLYPSNVNLTY-----TSDNTKAIESYNMSDNA-NKDEKFPDLSLA F 95
QY    239 KAEDLWGTDSDFMGFYQQOSHWOIFNGKNRPFRVHDIOPEIPLTPY-VSDLPMGKV 297
     + : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db    96 PLMRCILDNSILIGASYTORSMWQLSNTGSAPEFRETVEPOLFLFGATDYSVGDW--TL 153
QY    298 RHLGGVAHHSGESAKILSRSWNRAYLMAAGMKRLTYMPRIWGRIFFRGSSQDDNDP 357
     + : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db    154 RDAEFGYHNQSGRDPPTSRSMNRILYSRLMAONGWLVEVKFWYI-----GDTSDKN 207
QY    358 ILDYGYGVDFRELYOLEKKSNIQTGYRYNPRSGLALDIIDYYPLPGKGISGYFOIGYG 417
     + | | | : : : | : : | : | : | : | : | : | : | : | : | : | : | : | :
Db    208 ITKYMGYQLRKIGYL-GEAVLSAKGOYNMNNGYGAELGSYPITKHVREFTYGYSIG 266
QY    418 OSLDIDYNEATSFCYGLMLND 438
     :| ||| : : : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db    267 ESLIDYDFNQTRVGGMVMLND 287

RESULT           5
ID   GUN_BACSI              STANDARD:                PRG:      800 AA.
AC   P06564;
RT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   ENDOGUCANASE PRECURSOR (EC 3.2.1.4) (ENDO-1,4-BETA-GLUCANASE)
DE   (ALKALINE CELLULOSE).
OS   Bacillus sp. (strain 1139).
OC   Bacteria; Firmicutes; Bacillius/clostridium group;
OC   Bacillus/staphylococcus group; Bacillus.
OX   NCBI_TaxID=1411;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87085443; PubMed=3098909;
RA   "Molecular cloning and nucleotide sequence of the alkaline cellulase gene from the alkalophillic Bacillus sp. strain 1139."
RL   J. Gen. Microbiol. 132:2329-2335(1986).
RC   -1 CAVALTICTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC LINKAGES IN CELLULOSE.
CC   -1 MISCELLANEOUS: ALKALOPHILIC BACILLUS SP., STRAIN 1139, PRODUCES ONE ALKALINE CELLULOSE (PH OPTIMUM 9.0). THIS STRAIN IS NOT A TRUE CELULOTYTIC MICRO-ORGANISM BECAUSE THE ENZYME IS UNABLE TO HYDROLYSE NATIVE CELLULOSE.
CC   -1 SIMILARITY: BELONGS TO CELLULOSE FAMILY A (FAMILY 5 OF GLYCOSYL HYDROLASES).
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation -
```

Query Match	Best Local Similarity	4.6%;	Score 109;	DB 1;	Length 800;
Matches	91;	Conservative	51;	Mismatches 150;	Indels 170;
					Gaps 24;
CC	the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@lsb-sib.ch).				
CC	EMBL; M15743; AAA22305.1; -				
DR	EMBL; D00066; BAA00045.1; -				
DR	PIR; A29003; A29003.				
DR	InterPro; IPR001547; Glyco_hydro_F5.				
CC	Pfam; PF00150; cellulase; 1.				
DR	PROSITE; PS00659; GLYCOSYL_HYDROL_F5_1.				
KW	Cellulose degradation; Hydrolase; Glycosidase; Signal.				
FT	SIGNAL	1	30	POTENTIAL.	
FT	CHAIN	31	800	ENDOGLUCANASE.	
FT	ACT_SITE	190	190	PROTON DONOR (BY SIMILARITY).	
FT	ACT_SITE	305	305	NUCLEOPHILE (BY SIMILARITY).	
SQ	SEQUENCE	800 AA;	88602 MW;	7C0AD7BBDAD55CFC	CNC64;
QY	30	PVAFDEVRSKND-----	LGODNELLIIGVQTSQASDTANPL-DEHEPELY-----	TT	78
Db	25	PPLAAEGNTRDNEKHLNDNVNRPSEAGALQAEVGGOMTLVDGEEKIQLRGSMTH			84
QY	79	ALENTKMLINGCAL-----	NODIMRLACY-DTLVHETPAVVIKTRSRIRDET-----		126
Db	85	GLQWPEELINDNAYKALANDMESNIRLAMYGEINGVSNPELLISRYIKGIDLAIEDNM			144
QY	127	-----WQT-----	-----IKGKPOVNYQETDP-----	IFLMGN	150
Db	145	YVIYDMHVAHPDPRDPVYAGAEDEFRIALYPNNPHTIYELANEPSSNNNGAGIPNN			204
QY	151	EKGMTTKDAKOLEFAAQFPTLSFPLDRN-----	NPLWSSRP-----	HANPY	196
Db	205	BEGMAVNR-----	EVADPIVEMLRQSGNADDDIIIVGSPNNSQRPDLADNPIDDHRTMY		259
QY	197	VLPF--NHGKPNRS--	PMTPEHARQFTPNFEFARPELFGVSVYVKAEDLMGT-----		247
Db	260	TVHFTTGSAASTESYPPETPNSEGVNWSNTRYA-----	LENGAVFATE--WGTSQANG		313
QY	248	-----DSDLEFGYTOQ-----	-----SHMOIFNGKNS-----	RPRVH-----	275
Db	314	DGGPFDEADVIEFLFNENISMAMWSLTN--	KNEVSGAFTPEELKSNMNTSLDPEPDQVW		372
QY	276	-----	DIQPELFTLPQVYSDLPW--	DCKVMITGMGAVHHSNGSS--	312
Db	373	VPEELSLGGEVVARIKGVNEP--	IDRTKTKYLVMEFNCTKGQFGV-----	NGDSP	423
QY	313	-----	AKLSRSMNRAYLMAAGEMKKLGTWPRIMGR		342
Db	424	VEDVYIENBAGALKSLGIDASNDVSEGNWYANARLSADGMCK			465
RESULT 6					
GUN_PAPO					
ID	GUN_PAPO	STANDARD;	PRT;	397 AA.	
AC	P2348;				
DT	01-NOV-1991 (Rel. 20, Created)				
DT	01-FEB-1996 (Rel. 33, Last sequence update)				
DT	15-JUL-1998 (Rel. 36, Last annotation update)				
DE	ENDOGLUCANASE (EC 3.2.1.4) (ENDO-1,4-BETA-GLUCANASE) (CELLULASE).				
OS	Paenibacillus polymyxa (Bacillus polymyxa).				
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;				
CC	Bacillus/Staphylococcus group; Paenibacillus.				
RN	NCBI_TaxID=1406;				
RP	SEQUENCE FROM N.A.				
RA	MEDLINE=90170877; Pubmed=2307659;				
RA	Baird S.D., Johnson D.A., Seliqy V.T.;				

*Molecular cloning, expression, and characterization of
RT endo-beta-1,4-glucanase genes from *Bacillus polymyxa* and *Bacillus*
RT *circulans*.";
RL J. Bacteriol. 172:1576-1586(1990).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
CC LINKAGES IN CELLULOSE.
CC -1- SIMILARITY: BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL
CC HYDROLASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M33791; AAA22631.1; -
DR PIR: A35136; A35136.
DR HSSP: P54583; 1ECE.
DR InterPro: IPR001547; Glyco_hydro_F5.
DR Pfam: PF00150; cellulase_1.
DR PROSITE: PS00659; GLYCOSYL_HYDROL_F5; 1.
KM Cellulose degradation: Hydrolase; Glycosidase.
FT ACT_SITE 194 194 PROTON DONOR (BY SIMILARITY).
FT ACT_SITE 317 317 NUCLEOPHILE (BY SIMILARITY).
SQ SEQUENCE 397 AA; 44357 MW; B9C2E802C04F0A2A CRC64;

Query Match 4.0%; Score 94.5; DB 1; Length 397;
Best Local Similarity 25.7%; Pred. No. 2.7;
Matches 46; Conservative 21; Mismatches 61; Indels 51; Gaps 10;

QY 301 GCGAVHNSGESAKLSRSMRAYLMAGMEKNTLVPRMGR-----IFKEG----- 347
DB 43 GNRIVDSGKEAFNGLNW-----FGLERPNTY-LHGLMSRSDMDLDQVKEGYNLIR 95
QY 348 -----SGSQPDD-----NPDIIYVGQVGRFLYOLEKNSNIGTV-----RYMRS 389
DB 96 LPSYNOLFDSRRPDSIDYKKNPDLY---GLNPQLIMDKLTERAGGICITIDRRHPS 152
QY 390 GKGALDLDYV--YPLGKISGVFOIFQGY-----GSLIDYHNEATSFVGLMLNDW 439
DB 153 G-GQSELMWTISQYFESRWMISDMKMLADRYKNNPVIYGADLHNEPHGASNGTGNASTDW 210

RESULT 7
PKSM_BACSU STANDARD; PRT; 4273 AA.
AC P40872: O31781;
DT 01-FEB-1995 (Rel. 31, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PUTATIVE POLYKETIDE SYNTHASE PKSM.
GN PKSM OR PKSY.
OS *Bacillus subtilis*.
OC Bacteria; Firmicutes; *Bacillus*/Clostridium group;
OC *Bacillus*/Staphylococcus group; *Bacillus*.
OX NCBI_TaxID=1423;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE OF 1-1763 FROM N.A.
RC STRAIN=168 / Pbl1424;
RA Tognoni A., Grandi G.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: POTENTIALLY INVOLVED IN SOME INTERMEDIATE STEPS FOR
CC THE SYNTHESIS OF A POLYKETIDE MOLECULE WHICH MAY BE INVOLVED IN
CC SECONDARY METABOLISM.
CC -1- COFACTOR: CONTAINS 4 COVALENTLY BOUND PHOSPHOPANTETHEINES

(POTENTIAL).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z99113; CAB3603.1; -
DR EMBL: Z35133; CAA84505.1; -
DR Subtilist; BG10931; PKSM.
DR InterPro: IPR000794; ketoacyl-synt.
DR InterPro: IPR001601; Meth-transf.
DR InterPro: IPR003880; Phosphopant_attach.
DR Pfam: PF00109; ketoacyl-synt; 1.
DR Pfam: PF00550; pp-binding; 1.
DR PROSITE: PS00012; PHOSPHOPANTETHEINE; 2.
DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 2.
DR TRANSFERASE; ACYLTRANSFERASE; Antibiotic biosynthesis; NADP;
KM Phosphopantetheine; Multifunctional enzyme; Repeat; Complete proteome.
FT DOMAIN 295 364 ACYL CARRIER (ACP) 1.
FT DOMAIN 396 834 BETA-KETOACYL SYNTHASE 1.
FT DOMAIN 2190 2238 ACYL CARRIER (ACP) 2.
FT DOMAIN 2322 2737 BETA-KETOACYL SYNTHASE 2.
FT DOMAIN 3532 3947 BETA-KETOACYL SYNTHASE 3.
FT DOMAIN 3410 3483 ACYL CARRIER (ACP) 3.
FT DOMAIN 4140 4209 ACYL CARRIER (ACP) 4.
FT BINDING 327 327 PHOSPHOPANTETHEINE (POTENTIAL).
FT BINDING 2222 2222 PHOSPHOPANTETHEINE (POTENTIAL).
FT ACT_SITE 2476 2476 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
FT ACT_SITE 3446 3446 PHOSPHOPANTETHEINE (POTENTIAL).
FT ACT_SITE 3690 3690 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
FT BINDING 4172 4172 PHOSPHOPANTETHEINE (POTENTIAL).
FT CONFLICT 103 103 E -> V (IN REF. 2).
FT CONFLICT 276 276 E -> E (IN REF. 2).
FT CONFLICT 289 289 T -> S (IN REF. 2).
SQ SEQUENCE 4273 AA; 477459 MW; 3BBFC1A250AEB5A CRC64;

Query Match 3.9%; Score 93; DB 1; Length 4273;
Best Local Similarity 19.2%; Pred. No. 97;
Matches 102; Conservative 79; Mismatches 163; Indels 188; Gaps 29;

QY 14 PCFALIAIOQAQAV-PMPVAFVDE-----VRSKN-----DLGQ 45
DB 3597 PLFQIOPKREKESMDPROPIFLEAMHTFEDAGYMGDRKSGGVYVGESEGYAHLTG 3656
QY 46 DNEILLIGVOSATOSASTDTANPLDEHEPELYTALENKTMLI---NCSALNQDITRLACY 102
DB 3657 DTDYINGTQNTATLSAR--IAVALDLKGPNNMALTAACSSGLVAIHQACSLAROG---DCE 3710
QY 103 DTLVHGETPRAVITKRSIRLDEITWOTIKGRPVVYQETDPIRLMENEKMLTKKAKO 162
DB 3711 MALAAGVTLN-----SHMSFALTRAEKLSRNGQCKYVDQDANG 3750
QY 163 L-----EYAAKQFPLSLSFDLRRNNTPLMSGRPNPMVYLPIFMHGRPN--RSPNTSHE 216
DB 3751 LVREAVAAVILKPLSKAIE-DKDH-----YGCIRKASVNVNDGNTNGTANP--PFSQ 3801
QY 217 ARO-----FTPNEFRAPLEKFOVS-----VKVKAEDLWGTSDLMFGYTOQSHMQ 262
DB 3802 ALELTENIYERKEINPLDIQYVMAHSTGSLNCPDLEVALTSVFSK-----YTKROFC 3854
QY 263 IFNCKNRPRVHYOQPEIFITOPVYSDLPWDGKVRNIGM-----GAVHSHNGES- 312
DB 3855 MIS--SIKPLIGHTEA-----SGTVALISMAMKNOIIPATHHCESENP 3898
QY 313 -----AKLSRSMRAYLMAGMEKNTLVMPRMGRIFKEG--SGSQ----- 351

Db 3699 YIPKESFVLCKENRSWK-----KN-OKPRM-GTISTGISTNAHAVIEYI 3946
 QY 352 PDDNPDLIDYGYGV-----RFLYOLENKSNI-SGVRNPNRSGKAL 394
 Db 3947 PDDPSTGRHSGSPQIFPISANONDRLODACRMIAVLEQHNHNSLPVATTLQVGRAM 4006
 QY 395 QL-----DYVPL-GKISGYEQ-IFQYQGSLLIDYNEH 427
 Db 4007 EARLAIVANQEQVLKRYKLEYEAMKNGVSGQGSRLYGTGTEGLEBODEA 4058
 RESULT 8
 SRK6_BRAOL STANDARD; PRT: 849 AA.
 AC 009092;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE PUTATIVE SERINE/THREONINE KINASE RECEPTOR PRECURSOR (EC 2.7.1.37)
 DE (S-RECEPTOR KINASE) (SRK).
 GN SRK6.
 OS Brassica oleracea (Cauliflower).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Brassica.
 OC NCBI_TaxID=3712;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV_S6S6; TISSUE=Stigma;
 RX MEDLINE=92020942; PubMed=1681543;
 RA Stein J.C., Howlett B., Boyes D.C., Nasrallah M.E.;
 RT "Molecular cloning of a putative receptor protein kinase gene encoded
 RL at the self-incompatibility locus of Brassica oleracea.";
 Proc. Natl. Acad. Sci. U.S.A. 88:8816-8820(1991).
 CC -1- FUNCTION: INVOLVED IN SPOROPHYTIC SELF-INCOMPATIBILITY SYSTEM
 (THE INABILITY OF FLOWERING PLANTS TO ACHIEVE SELF-
 CC FERTILIZATION), PROBABLY ACTING IN COMBINATION WITH S-LOCUS-
 CC SPECIFIC GLYCOPROTEINS. INTERACTION WITH A LIGAND IN THE
 CC EXTRACELLULAR DOMAIN TRIGGERS THE PROTEIN KINASE ACTIVITY OF THE
 CC CYTOSOLIC DOMAIN.
 CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN -> ADP + A PHOSPHOPROTEIN.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY IN THE PISTIL AND ANTHER.
 CC -1- POLYMORPHISM: THERE ARE A NUMBER OF DIFFERENT S ALLELES IN
 CC B. OLERACEA, POSSIBLY PROVIDING THE RECOGNITION SPECIFICITY.
 CC -1- SIMILARITY: THE EXTRACELLULAR DOMAIN IS SIMILAR TO S-LOCUS
 CC GLYCOPROTEINS OF BRASSICA, WHILE THE INTRACELLULAR DOMAIN IS
 CC A SER/THR-PROTEIN KINASE RELATED TO RAF KINASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M76647; AAA33000.1; ALT_TERM.
 DR HSSP: P1362; JGI.
 DR InterPro: IPR001480; B_lectin.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR003609; Pan_app.
 DR InterPro: IPR002290; Ser_thr_kin_actsite.
 DR InterPro: IPR000858; Slocus_glycop.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF00954; S_locus_glycop; 1.
 DR SMART: SM00473; B_lectin; 1.
 DR SMART: SM00473; PAN_AP; 1.
 DR SMART: SM00221; STYC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
 KW transferase; Serine/threonine-protein kinase; signal; ATP-binding;

KW Transmembrane; Receptor; Glycoprotein; Self-incompatibility.
 FT SIGNAL 1 32
 FT PUTATIVE SERINE/THREONINE KINASE
 FT CHAIN 33 849
 FT
 FT DOMAIN 33 446
 FT TRANSMEM 447 466
 FT DOMAIN 467 849
 FT DOMAIN 528 779
 FT NP_BIND 534 542
 FT BINDING 556 556
 FT ACT_SITE 653 653
 FT CARBOHYD 47 47
 FT CARBOHYD 120 120
 FT CARBOHYD 196 196
 FT CARBOHYD 260 260
 FT CARBOHYD 314 314
 FT CARBOHYD 389 389
 FT CARBOHYD 442 442
 SQ SEQUENCE 849 AA; 97231 MW; 7E156059BDE4370 CRC64;
 Query Match 3.9%; Score 92.5; DB 1; Length 849;
 Best Local Similarity 21.9%; Pred. No. 11;
 Matches 78; Conservative 50; Mismatches 107; Indels 121; Gaps 24;
 QY 105 LVHGETPAVVKTKRSIRLDETITQTKGPOVYQETDPIFLMG-----NEK--GMLTR 157
 Db 24 LHPALSTIYNT-----LSTESTLTSNKTLV-----SPQISEVGFPRNSRWYLGMYWK 75
 QY 158 K-DAKOLEYAKQFTPLSLSPD-----LDRNNTPLMSSRPHNPMYVLPFMNGK 205
 Db 76 KVSDETYVWVNRDPLNSAIGTLKISGNNLVLDHSMKPVMTN-----LTRGN 125
 QY 206 PKRSP-----NTPSHEARQPT-----PNEFRAPELKFQSVYKKAEDLM 245
 Db 126 -ERSVYVAFELANGFNWRDSSNNDASEYLWQSFDPIDTLPEKKLGYNLKT----- 177
 QY 246 GTDSULMEGYTQOSHMOJFNGKNSRPFYVDYQ-----PELFLNQ--PYVSDLPWDGK 296
 Db 178 GLNRFL-----TSWRSSDDPSSCNF---SYKLETFQSLPEFYLSRENPMPHRSQPMWG- 226
 QY 297 VNMIGMAVH-----HNGESAKLSRSMNAY-----LMAGMEKNLTIVMP--RI 339
 Db 227 IFFSIPEDOKLSYVWYNEIENNEEVAITYFRMTNNSFSRLTISEGTFORLTWPSIRI 286
 QY 340 WQRIKESGSGOPDNPDLIDY-----YGYGVRFLEYOLENKSNIQVYR-VNPRS 389
 Db 287 WNRFW-----SSPVD-PQCDTYIMCGPYAYCDY-----NTPVNCNCIQGPNPRN 329
 RESULT 9
 YL70_ARCFU STANDARD; PRT: 324 AA.
 ID YL70_ARCFU
 AC 028112;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOTHETICAL PROTEIN AF2170.
 GN AF2170.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
 OC Archaeoglobus.
 OC NCBI_TaxID=2234;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.R., Peterson J.D.,
 RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrides N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirtness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeill L.K., Badger J.H., Glodek A., Zhou L.,

RA Overbeek R., Gocayne J.D., Welden J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sadow P.M., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon *Archaeoglobus fulgidus*.";
RL Nature 390:364-370(1997).
CC -1- SIMILARITY: CONTAINS 2 KELCH REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000955; AAB89093.1; .
DR TIGR: AF2170; .
DR InterPro: IPR001798; Kelch.
DR Pfam: PF01344; Kelch; 2.
KW Hypothetical protein; Repeat; Complete proteome.
FT REPEAT 229 276 KELCH 1.
FT REPEAT 277 323 KELCH 2.
SQ SEQUENCE 324 AA; 36025 MM; C445388CFEB96E45 CRC64;

Query Match 3.9%; Score 92; DB 1; Length 324;
Best Local Similarity 20.1%; Pred. No. 3.2;
Matches 54; Conservative 44; Mismatches 72; Indels 98; Gaps 18;

OY 224 EFRAPLEKFOYSKVKAAED-----LKGTSDFMGFYTOOSHWQIF-NGKNSRPFVH 275
DB 63 EFRPN-RLRI-LLSSEDFSGREAAVWQDEELIFGTT-----VFENKYSPTDQIL 112
OY 276 DYQPEIRLTQPVYSDLP-----W-DGKVRMI-----GGAHVHSGESAKLSRS 318
DB 113 SFNPKLERLVANSLPHRPSDVAANWGDNRVYIFLNNSCEYAYAFPSSESRAKLDS 172
OY 319 -----W-NRAYL-----MAGEMKMLVMPRIWGRIFREGSGSQ 351
DB 173 CPHEPGCVHSAVYWGKAYFCGEGVASFDPMGFKW--IATLDYRWVAARVAADG-- 228
OY 352 PDDNPDLIDYGYGDVFLVLEKNSISGT-----VRNPSSGKALDLYVPLGKC-- 405
DB 229 -----YFAIGSSGIAETKEIRIRNPXTGE-LCEMRKLPVARGQA 270
OY 406 --ISG-YFOIF-----OGYGSLIDYNH 425
DB 271 VAVGGEYIYIFGTYTKDGYANEIIRYD 298

RESULT 10
SCKL_SCHPO STANDARD; PRT: 696 AA.
AC PF0530; O9UTP3;
DT 01-OCT-1996 (Rel. 34, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SRINE/THREONINE-PROTEIN KINASE SCK1 (EC 2.7.1.37).
GN SCK1 OR SPAC1B9.02C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96120227; PubMed=7498728;
RX Jin M., Fujita M., Culley B., Apollinaro E., Yamamoto M.,
RA Maundrell K., Hoffman C.;
RT "sck1, a high copy number suppressor of defects in the CAMP-dependent

RT protein kinase pathway in fission yeast, encodes a protein homologous
RT to the Saccharomyces cerevisiae Sch9 kinase.";
RL Genetics 140:457-467(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA McQuigley R.C., Rajandream M.A., Barrell B.G., Whitehead S.,
RA Churcher C.M.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN = ADP + A PHOSPHOPROTEIN.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CAMP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D38108; BAA07286.1; .
DR EMBL: AL109951; CAB53053.1; .
DR HSSP: P05132; ICTP.
DR InterPro: IPR000008; C2.
DR InterPro: IPR000719; Euk_Pkinase.
DR InterPro: IPR000961; Pkinase_C.
DR InterPro: IPR002290; Ser_Thr_kin_actsite.
DR Pfam: PF00069; Pkinase_1.
DR Pfam: PF00433; Pkinase_C; 2.
DR SMART: SM00239; C2; 1.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00133; S_TKc_X; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP_1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST_1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase: Serine/threonine-protein kinase; ATP-binding; CAMP.
FT DOMAIN 302 563 PROTEIN KINASE.
FT NP_BIND 308 316 ATP (BY SIMILARITY).
FT BINDING 331 331 ATP (BY SIMILARITY).
FT ACT_SITE 428 428 BY SIMILARITY.
FT CONFLICT 199 199 A -> R (IN REF. 1).
SQ SEQUENCE 696 AA; 78594 MM; A7B05F5E24A2A7 CRC64;

Query Match 3.9%; Score 92; DB 1; Length 696;
Best Local Similarity 19.2%; Pred. No. 9.3;
Matches 87; Conservative 66; Mismatches 163; Indels 138; Gaps 19;

OY 16 FAIIAIOQAQ-VPRPVAFVD-----EVRSKN-----DLGQDNELLIG----- 52
DB 161 YAVITFEKTYQVWPPRPFIDGIPISIPSKNRPPLAGSASSGSLHSELMADYRCPHW 220
OY 53 ----VQATQASSTDTANPLDHEPELTYTALNKTMLINCASALNODIMRLACYDTLVHG 108
DB 221 DPEIVFDVTKKKSQMWVYVDKEDKFLGSVKITPIFLH-----EVQGEWYKL----- 270
OY 109 ETPAVIKTKRSIRLDETITQTIK-----GKPVVYQETDPIFLM 148
DB 271 EPLDITKLEGEIKETIYEHENHVRGPEDFTLRLIGKSTFGQVYLVRKNDTRIRYAM 330
OY 149 GNEKGMILTKRAKOLEYAAKQFTPLSLFDDRRNTPMLSSRPNHP----- 194
DB 331 -----KTSKSKLIVRKKEVT-----HTLGERIILVTRTSLDESPTVGLKFSFQTASD 377
OY 195 MYVLPIFMHGKPNRSPNTPNSHEARQFTNPFAPR--LKFOVSYVKAEDLMGTSDLM 252
DB 378 LYLTIDVMSG-----GELFMHLOHGBRFRQRAKIFYIAELVLALENH----- 419
OY 253 FGYYTQSHMOIFNGKNSRPFVRYDQREIFLTQPVYSDLPMDGKVRMTGMCAVHHSNGES 312
DB 420 -----HKHDIIY-----RDLKPEINIT-----LDADGHIALCDFGISKANLSAN 457

QY 313 AKLSRSMNRAYLMAAG---MEMKNLTVMPTIRWG---RIEKEGSGSQPDNDPILDIY---G 363
 DB 458 ATTNFCGTTEETLAEVLEEDKGYTKQVDFMSIGLVLFEMCCGMSPFYAPDQOMYRNA 517
 QY 364 YGDVRE---LYOLEKNSINSGVRYRNPSSKCAL 394
 DB 518 FGKVRFPKGVLSSEGRSFPVRLNPNHRLGAV 551

RESULT 11

YJHO_YEAST STANDARD: PRT: 888 AA.
 AC P40361;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE HYPOTHEMETICAL 104.3 KDA. PROTEIN IN SMC3-KRP18 INTERGENIC REGION.
 GN YJL070C OR J1095 OR HRD888.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Saccharomycetales; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 NX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C;
 RA MEDLINE=95282514; PubMed=7762302;
 RX Vandenbol M., Durand P., Dion C., Portetelle D., Hliger F.;
 RT "Sequence of a 17.1 kb DNA fragment from chromosome X of
 RT Saccharomyces cerevisiae includes the mitochondrial ribosomal protein
 RT L8.";
 RL Yeast 11:57-60(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C;
 RA Sor F.J.;

Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE ADENOSINE AND AMP DEAMINASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: Z34288; CAA84052.1; -
 DR EMBL: Z49345; CAA89362.1; -
 DR EMBL: Z49344; CAA89361.1; -
 DR EMBL: X88851; CAA61309.1; -
 DR PIR: S47120; S47120.
 DR SGD: S0003606; YJL070C.
 DR InterPro: IPR001365; A.deaminase.
 DR Pfam: PF00962; A.deaminase.1.
 KM Hypothetical protein; Hydrolase.
 SQ SEQUENCE 888 AA; 104263 MW; DB31A808622411AD CRC64;

Query Match 3.9%; Score 92; DB 1; Length 888;

Best Local Similarity 22.1%; Pred. No. 13;
 Matches 66; Conservative 36; Mismatches 126; Indels 70; Gaps 15;

QY 128 QTIKGRPOVVOE---TDPFLMGNEKMLTKKDA-----KOLEYAAKQFPLSL 176
 DB 2 QAVERRPRLPDEQNSVTKPNETKNKEARVLENDGDVSPVLKQKEISVDMDMISLP 61
 QY 177 FDLDRN---NTPLM---SSPHNPMYVLPFTFMGKPNR-----SPVTPSH-----EAPRF 220
 DB 62 TEEDKQMLVSGPMFPLDEDEENKIDPLPSVSHHYGSEDSFVSSTPNSLKTGEETKDL 121
 QY 221 TNERRAPLTKQVSVKVAADLMG-----TDSDLWFGYTQ---QSHWQIFNKKNSRPR 273
 DB 122 FLPNPF---ELVSQMKRKYTAASKQDISINIKNDTEKWFLLPKPLKFWREDDK-----R 173

QY 274 VHDYQPELFTOPYVSDLPWDMGKVRMIGAGVAHNSGESAKLSRSMNRAYLMAHEMKNL 333
 DB 174 FQD-----PDSDLNDDGDSGTGAATPHRHGYYPYFTDHYXXYTKSGLKGN 224
 QY 334 TWMPTIRGRIEKEGSGSQPDNDPILDIYGYGDVRLYOLEKNSINSGVRYRNPSSK 391
 DB 225 IKVP-YTGEYF-----DLEDY---KKOYIYHLSNQENTQNP--SPYSSK 263

RESULT 12

SACB_STRSL STANDARD: PRT: 969 AA.
 ID SACB_STRSL
 AC Q55242;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE LEVANSUCRASE PRECURSOR (EC 2.4.1.10) (BETA-D-FRUCTOFURANOSYL
 DE TRANSFERASE) (SUCROSE 6-FRUCTOSYL TRANSFERASE).
 GN FTF.
 OS Streptococcus salivarius.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 NX NCBI_TaxID=1304;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 25975;
 RX MEDLINE=9332332; PubMed=8331080;
 RA Ratham C., Giffard P.M., Jacques N.A.;

"The cell-bound fructosyltransferase of Streptococcus salivarius: the
 RT carboxyl terminus specifies attachment in a Streptococcus gordoni
 RT model system."
 RL J. Bacteriol. 175:4520-4527(1993).
 CC -1- CATALYTIC ACTIVITY: SUCROSE + (2, 6-BETA-D-FRUCTOSYL)(N) -
 CC GLUCOSE + (2, 6-BETA-D-FRUCTOSYL)(N+1) (OTHER SUGARS CAN
 CC ALSO ACT AS FRUCTOSYL ACCEPTORS).
 CC -1- SUBCELLULAR LOCATION: SECRETED. CELL-WALL BOUND.
 CC -1- SIMILARITY: BELONGS TO FAMILY 68 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: L08445; AAA71925.1; -
 DR InterPro: IPR003469; Glyco_hydro.68.
 DR Pfam: PF02435; Glyco_hydro.68; 1.
 DR Transferrase; Glycosyltransferase; Signal; Cell wall.
 FT SIGNAL 1
 FT CHAIN 1 969 LEVANSUCRASE.
 SQ SEQUENCE 969 AA; 103983 MW; D389B5B32AC7735A CRC64;

Query Match 3.9%; Score 91.5; DB 1; Length 969;

Best Local Similarity 19.9%; Pred. No. 16;
 Matches 85; Conservative 54; Mismatches 161; Indels 127; Gaps 20;

QY 53 VQSATOSATIDTPANPLDEHEPELYTALENKTYMLNCSALNDIMLA---CYDPLVNG 108
 DB 155 VEAPITSTASSEAA---DTH-TEVDLKVSSNSANANMLSKLNGRIKISIVEENMTSDIY-A 209
 QY 109 ETPAVIKTKRSIRLDEITQITGK-POVVOETTPFLMGNEKMLTKKDAKOLEYAA 167
 DB 210 LTFEELKALNKVDFSD---DAIKGTSTSLTRYLAKDIY-----ASFLLKDSKLAVPF 259
 QY 168 KQTPPLSL-SFD-----LDRNTPLMSSRPNPM-----YLPDIFMHGKPNRSP 210
 DB 260 KADTIINMPAFNVDQATKKEIDVDWDMVPQDAKSGVSNMNGYQLVISMAGAPKNKS 319
 QY 211 NPFSHARQFTPNERAPLTKQVSVKVAADLMGCTSDSLWFGYTGQSHWQIFNKKNSR 270


```

Db      320 NHILLYRKYGNDP-----THM-----KNAG 341
Qy      271 PEFVHYOPEIFLTOPYSDLPMDGKRYMIGCAV--HHSNGESAKLSRWMNR---AYLM 325
Db      342 P-----IFGYNMLEDDOQWGSATVNSDGSIDLYTKNDTSGKLMWOLASATLN 392
Qy      326 AGMEMKNTLVMPRIWGRIFKEGSG-----SOP-----DDNPDILDYGYGD--- 366
Db      393 LAVEENDEVIKSVENDHILFEGDNYHYQSTPKMSTFPDDNDHNDGNDPDRDNCCLRPHI 452
Qy      367 -----VREL-----YOLENK-----SNIGTVYVNRSGKALQLDYVYPLGKIS 407
Db      453 IEDNGSRVILFESNTGDENQGEKQIYKMSNYGDDAFNLKSPNLNHNKHLNLASMAN 512
Qy      408 GYFOIFQ 414
Db      513 GSIGILK 519

```

```

RESULT 13
CATA_DICDI STANDARD; PRT; 496 AA.
ID CATA_DICDI
AC 077229;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CATALASE (EC 1.11.1.6).
GN CATA OR CAT.
OS Dictyostellium discoideum (Slime mold).
OC Eukaryota; Metazoa; Dictyostellida; Dictyostellium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RA Foote C., Alexander H., Alexander S.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
PEROXIDE.
CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC -1- COFACTOR: HEME GROUP (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AF090443; AAC36743.1;
DR HSSP: P00432; 7CAT.
DR DictyDb: DP027272; cata.
DR InterPro: IPR002226; Catalase.
DR Pfam: PF00199; catalase.1.
DR PRINTS: PR00067; CATALASE.
DR ProDom: PD000510; Catalase.1.
DR PROSITE: PS00437; CATALASE.1; 1.
DR PROSITE: PS00438; CATALASE.2; 1.
KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Peroxisome.
FT ACT_SITE 54 54 BY SIMILARITY.
FT ACT_SITE 128 128 BY SIMILARITY.
FT BINDING 338 338 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT SITE 494 496 MICROBODY TARGETING SIGNAL (POTENTIAL).
SQ SEQUENCE 496 AA; 55683 MW; 68331888BFAFDZE6 CRC64;

```

Query Match 3.9%; Score 91; DB 1; Length 496;
Best local Similarity 23.5%; Pred. No. 7;

```

Matches 56; Conservative 32; Mismatches 94; Indels 56; Gaps 12;
Qy      61 STDANPLDEHEPELTYTALENKTMILNCALNDIMRLACYPVLGEPFAVYKTRRSI 120
Db      7 TSSGSPIDNN-----LNSMTAGVNGPILLIDFTLI---DKLAHFDR----- 46
Qy      121 RLDETITWQTIKGRQVYVOETTD-----PIF-----LMGNEKML-T 156
Db      47 RIERRVVHA-KGAGAHGFEYTSDDVKKCKAKFLNKVGRKTPFTFSTVYGGEGSSDS 105
Qy      157 KDAKQLEVAKQFTPLSLFDDLRNNTPLMSSRPHPMYVLPIFMHKKRSPSTPSHE 216
Db      106 ERDPR--GFAVKFYTEEG-NPDWVGNTPVFFTRIDPSK---FPDFITQ-KRNPQTCKD 158
Qy      217 AROFTPNEFRAPELKFOYSYKVAEDLKGTDSDLMFGTYQOSHWOIFNGKNSRPFVY 274
Db      159 PNMEWDFLGOTPESTHQVSILFSDR---GTPKSYRMHGFSSHTLKEVNAQGRPYWY 212

```

```

RESULT 14
A2MG_HUMAN STANDARD; PRT; 1474 AA.
ID A2MG_HUMAN
AC P01023;
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ALPHA-2-MACROGLOBULIN PRECURSOR (ALPHA-2-M).
GN A2M.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85190481; PubMed=2581245;
RA Kan C.-C., Solomon E., Belt K.T., Chain A.C., Hlorns L.R., Fey G.H.;
RT "Nucleotide sequence of cDNA encoding human alpha 2-macroglobulin and
assignment of the chromosomal locus.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2282-2286(1985).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT HIS-704.
RX MEDLINE=92246939; PubMed=1374237;
RA Mathijs G., Devriendt K., Cassiman J.-J., van den Berghe H.,
RL Blochem. Biophys. Res. Commun. 184:596-603(1992).
RN [3]
RP SEQUENCE OF 24-1474.
RX MEDLINE=84239807; PubMed=6203908;
RA Sottup-Jensen L., Stepanik T.M., Kristensen T., Wierzbicki D.M.,
RA Jones C.M., Loenblad P.B., Magnusson S., Petersen T.E.;
RT "Primary structure of human alpha 2-macroglobulin. V. The complete
structure.";
RL J. Biol. Chem. 259:8318-8327(1984).
RN [4]
RP ERRATUM.
RA Sottup-Jensen L., Stepanik T.M., Kristensen T., Wierzbicki D.M.,
RA Jones C.M., Loenblad P.B., Magnusson S., Petersen T.E.;
RL J. Biol. Chem. 260:6500-6500(1985).
RN [5]
RP INHIBITORY SITE.
RX MEDLINE=84030513; PubMed=6195065;
RA Virca G.D., Salvesen G.S., Travis J.;
RT "Human neutrophil elastase and cathepsin G cleavage sites in the bait
region of alpha 2-macroglobulin. Proposed structural limits of the
bait region.";
RL Hoppe-Seyler's Z. Physiol. Chem. 364:1297-1302(1983).
RN [6]
RP INHIBITORY SITE.
RX MEDLINE=81212827; PubMed=6165619;
RA Sottup-Jensen L., Loenblad P.B., Stepanik T.M., Petersen T.E.,
RA Magnusson S., Joernvall H.;
RT "Primary structure of the 'bait' region for proteinases in alpha 2-

```

RT macroglobulin. Nature of the complex.";
RN FEBS Lett. 127:167-173(1981).
RN [7]
RN INHIBITORY SITE.
RX MEDLINE=81255805; PubMed=6167263;
RA Hall P.K., Mellies L.P., Travis J., Roberts R.C.;
RT "Protolytic cleavage sites on alpha 2-macroglobulin resulting in
RT proteinase binding are different for trypsin and Staphylococcus
RN aureus V-8 proteinase.";
RN Biochem. Biophys. Res. Commun. 100:8-16(1981).
RN [8]
RN INHIBITORY SITE.
RX MEDLINE=82095610; PubMed=6172288;
RA Mortensen S.B., Sotttrup-Jensen L., Hansen H.F., Petersen T.E.,
RA Magnusson S.;
RT "Primary and secondary cleavage sites in the bait region of alpha 2-
RT macroglobulin.";
RN FEBS Lett. 135:295-300(1981).
RN [9]
RN SEQUENCE OF 672-747.
RX MEDLINE=90242963; PubMed=1692292;
RA Mayrén P., Devriendt K., van den Bergh H., Cassiman J.J.;
RT "A genetic polymorphism in a functional domain of human pregnancy
RT zone protein: the bait region. Genomic structure of the bait domains
RT of human pregnancy zone protein and alpha 2 macroglobulin.";
RN FEBS Lett. 262:349-352(1990).
RN [10]
RN STRUCTURE BY NMR OF 1337-1474.
RX MEDLINE=99081553; PubMed=9865555;
RA Huang W., Dolmer K., Liao X., Gettins P.G.W.;
RT "Localization of basic residues required for receptor binding to the
RT single alpha-helix of the receptor binding domain of human alpha2-
RT macroglobulin.";
RN Protein Sci. 7:2602-2612(1998).
RN [11]
RN VARIANTS TYR-972 AND ILE-1000.
RX MEDLINE=92128897; PubMed=1370808;
RA Poller W., Feder J.-P., Klobbeck G., Olek K.;
RT "Cloning of the human alpha 2-macroglobulin gene and detection of
RT mutations in two functional domains: the bait region and the
RT thiolester site.";
RN Hum. Genet. 88:313-319(1992).
CC -1- FUNCTION: IS ABLE TO INHIBIT ALL FOUR CLASSES OF PROTEINASES BY A
CC UNIQUE "TRAPPING" MECHANISM. THIS PROTEIN HAS A PEPTIDE STRETCH,
CC CALLED THE "BAIT REGION" WHICH CONTAINS SPECIFIC CLEAVAGE SITES
CC FOR DIFFERENT PROTEINASES. WHEN A PROTEINASE CLEAVES THE BAIT
CC REGION, A CONFORMATIONAL CHANGE IS INDUCED IN THE PROTEIN WHICH
CC TRAPS THE PROTEINASE. THE ENTRAPPED ENZYME REMAINS ACTIVE AGAINST
CC LOW MOLECULAR WEIGHT SUBSTRATES (ACTIVITY AGAINST HIGH MOLECULAR
CC WEIGHT SUBSTRATES IS GREATLY REDUCED). FOLLOWING CLEAVAGE IN THE
CC BAIT REGION A THIOLESTER BOND IS HYDROLYZED AND MEDIATES THE
CC COVALENT BINDING OF THE PROTEIN TO THE PROTEINASE.
CC -1- SUBUNIT: HOMOTETRAMER, WHICH CONSISTS OF TWO PAIRS OF DISULFIDE-
CC LINKED CHAINS.
CC -1- TISSUE SPECIFICITY: PLASMA.
CC -1- MISCELLANEOUS: CONTRARY TO THE RAT PROTEIN, WHICH IS AN ACUTE
CC PHASE PROTEIN, THIS PROTEIN IS ALWAYS PRESENT AT HIGH LEVELS IN
CC CIRCULATION.
CC -1- SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,
CC INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation on
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M11313; AA51551.1; -;
DR PIR; A01256; MAHU.
DR PIR; S09107; S09107.
DR PDB; 1BV8; 30-SEP-98.

DR	SWISS-2DPAGE: P01023; HUMAN.
DR	MIM: 103950; -.
DR	InterPro: IPR002890; A2M_N.
DR	InterPro: IPR001599; Alpha_2_macroglbIn.
DR	Pfam: PF00207; A2M; 1.
DR	Pfam: PF01835; A2M_N; 1.
DR	PROSITE: PS00477; ALPHA_2_MACROGLOBULIN; 1.
KW	Serine protease inhibitor; Glycoprotein; Plasma; Bait region; Signal;
KW	3D-structure; Polymorphism.
FT	SIGNAL 1 23
FT	CHAIN 24 1474
FT	SITE 693 694
FT	DOMAIN 690 728
FT	SITE 704 709
FT	SITE 719 723
FT	SITE 730 735
FT	DISULFID 48 86
FT	DISULFID 251 299
FT	DISULFID 269 287
FT	DISULFID 278 431
FT	DISULFID 470 470
FT	DISULFID 563 563
FT	DISULFID 595 771
FT	DISULFID 642 689
FT	DISULFID 821 849
FT	DISULFID 847 883
FT	DISULFID 921 1321
FT	DISULFID 1079 1127
FT	DISULFID 1352 1467
FT	THIOLEST 972 975
FT	CARBOHYD 991 991
FT	CARBOHYD 55 55
FT	CARBOHYD 70 70
FT	CARBOHYD 247 247
FT	CARBOHYD 396 396
FT	CARBOHYD 410 410
FT	CARBOHYD 869 869
FT	CARBOHYD 1424 1424
FT	VARIANT 704 704
FT	VARIANT 972 972
FT	C -> Y (PROBABLY INTERFERES WITH THE ACTIVITY).
FT	/FTID-VAR_000012.
FT	VARIANT 972 972
FT	C -> Y (PROBABLY INTERFERES WITH THE ACTIVITY).
FT	/FTID-VAR_000013.
FT	VARIANT 1000 1000
FT	V -> I.
FT	/FTID-VAR_000014.
FT	CONFLICT 63 63
FT	CONFLICT 563 563
FT	MISSING (IN REF. 3).
FT	C -> E (IN REF. 3).
FT	SEQUENCE 1474 AA; 163278 MW; F1DS1742791GCFE CRC64;
SO	
Query Match	3.9%; Score 91; DB 1; Length 1474;
Best Local Similarity	21.0%; Pred. No. 32;
Matches 100; Conservative 72; Mismatches 157; Indels 148; Gaps 27;	
QY	4 SLSLTLSLPCFALITAGQAQAVNPVAFVDEVSKND--LGQDNELLIGVQSAQTQAS 61
DB	10 SLVLLLVLLPLPDASGKRPQYMWLVPSLHETTEKGCYLLSYLNE-----TVTSAS 63
QY	62 TDTANPLDEHEBELYTTALENKTMLINC-----SALQDIMRLACYDTLVHG----- 108
DB	64 LESV-----RGNSSLFTD-LEANDVLHCVAFAVPSSSSNEEVMFLP---VQYKGPQDEK 115
QY	109 -ETPAVIKTKRS---IRLDEITW---QITKGKPPVYYQGTTPD-----IFLMNGEKG 154
DB	116 KRTTVMKKNEDLVFQDTKSIYKPGQYKFR-VVSMDENFHPLNELPLVLYIDPQGNR 174
QY	155 LKKKAKOLEYNAKQTPPLSLSFDDIRNNTPLWSSRPHPMVLPLPFMHGKRNRSPTPS 214
DB	175 INQWQSFOLEGGLKQF-----SFPL-----SSPEFGSTKYVV-----QKSSGGR 216
QY	215 HEARQFTPEHFRAPELKQGVSV-----KVK 239

```

Db 217 H-----PFTVEEVLKRFKFEVQVTPVKITITILEEMNVSVCGLTYTGKRPVGHVTVSICKKYS 273
OY 240 AADLNGTSDLMFQYTOOSHMOIFNGK-NSRPFVHDIQPELFLQPVYSIDLPMGDKVR 298
Db 274 DASDCGEBS-----DAFCERFSGQLNSHGCIFYQVTKVFOLKREKEYEMKLTLEAQ 325
OY 299 MIGHGAVHNSNG-ESAKLSNSNR-----AYLMAGMEKNLTVMPRIWGRI-FKRGSG 349
Db 326 IQEETVVELTGRKSSSEITRTITKLSFVKVDSHFROGIFP-----FGQVRLVDGKG 376
OY 350 SOPDDNDPIL-----DYGG-----YGDVRFLYOLENKSNIQST-----VRYNPRS 389
Db 377 -VPIPKVIFIRGNEANYSNATDEHGLVQFSI-----NTINWGTSLTVAVNPKDRS 429

```

RESULT 15

```

INVO_MOUSE STANDARD; PRT; 467 AA.
ID INVO_MOUSE
AC P48997;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE INVOLUCRIN.
GN IVL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIH SWISS;
RX MEDLINE=94104476; PubMed=8277848;
RA DJan P., Phillips M., Easley K., Huang E., Simon M., Rice R.H.,
RA Green H.;
RT "The involucrin genes of the mouse and the rat: study of their shared
RT repeats."
RL Mol. Biol. Evol. 10:1136-1149(1993).
CC -!- FUNCTION: INVOLUCRIN IS A KERATINOCYTE PROTEIN THAT FIRST APPEARS
CC IN THE CELL CYTOSOL, BUT ULTIMATELY BECOMES CROSS-LINKED TO
CC MEMBRANE PROTEINS BY TRANSGLUTAMINASE. ALL THAT RESULTS IN THE
CC FORMATION OF AN INSOLUBLE ENVELOPE BENEATH THE PLASMA MEMBRANE.
CC -!- TISSUE SPECIFICITY: PRESENT IN KERATINOCYTES OF EPIDERMIS AND
CC OTHER STRATIFIED SQUAMOUS EPITHELIA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L28819; AAA39330.1; -.
DR MGD: MGI:96626; IYL.
DR InterPro: IPR002360; Involutrin.
DR PROSITE: PS00795; INVOLUCRIN; 1.
KW Keratinocyte; Repeat.
SQ SEQUENCE 467 AA; 54919 MW; 603E1E51B435737D CRC64;

```

Query Match 3 8%; Score 90.5; DB 1; Length 467;

Best Local Similarity 19.8%; Pred. No. 7.1; 128; Indels 79; Gaps 14;

```

OY 134 PQVYV-----QETTPDIFLGM-----NEKGLTKKDAQOLEFAAKOFTPLSLSFDRN 182
Db 175 PQELHLRQHQEKQDPELHNGOQKTPREBKLIPEGKQDELHNGRQEPQDEQLHNGOK 234
OY 183 NTPLMSSRPANPMYVLPIDENHGKPNRSPNTPSHEARQFTPNFEFAPDLKFOVSVKYKAAE 242
Db 235 Q-----KQKLHEPELQKQOQHOKPS--EPQLPLGKQOQESPE---PELPLGKQOQESPE 285
OY 243 DLMGTSDLMFGTQOSHNGIFNGKNSRPRVNDYQPEIFLQPVYSIDLPMGDKVRKIGM 302

```

```

Db 286 -----PELQLGKQOQSHFEDMAGDQKEREKRLH-----KPEYLIRKQOYQESP--DPE---LSL 333
OY 303 GAVHNSNGESAKLSRSWNRAYLMAGMEKNLTVMPRIWGRI-FKRGSGSOPDDNDPILDY 362
Db 334 GKQOHQCEPELQ-----LEEKQHKRPEPELHLGKQOQESHFEDMAED----- 377
OY 363 GYGDVRFLYOLENKSNIQSTVRYNPRSNGALQLDIYYPILGKISGYFQIFQYGGSL-- 420
Db 378 -----LEEKQKLG-----EPQLHLGKQOQOQIEREG-----YQGPKSIGOSLQ 416
OY 421 -----IDYNH 425
Db 417 EKASREQQLDYSH 429

```

Search completed: November 30, 2001, 14:27:03
Job time: 550 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:26:25 ; Search time 77.9 Seconds

(Without alignments)
829,941 Million cell updates/sec

Title: US-09-787-083-4

Perfect score: 2360
Sequence: 1 MKVSLSTLTSLPCFALTA.....YNHEATSRGVLMLNDMGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_17:*

1: sp.archaea:*\n2: sp.bacteria:*\n3: sp.fungi:*\n4: sp.human:*\n5: sp.invertebrate:*\n6: sp.mammal:*\n7: sp.mhc:*\n8: sp.organelle:*\n9: sp.phage:*\n10: sp.plant:*\n11: sp rodent:*\n12: sp.virus:*\n13: sp.vertebrate:*\n14: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	815	34.5	382	2	09K0U7
2	815	34.5	409	2	09JY21
3	373.5	15.8	306	2	09C122
4	366	15.5	329	2	09PMU8
5	360.5	15.3	292	2	09SID7
6	360	15.3	289	2	09L6N9
7	356.5	15.1	292	2	09Z4N8
8	342.5	14.5	297	2	03J349
9	246.5	10.4	355	2	02S241
10	240.5	10.2	355	2	09ZLX5
11	121	5.1	278	2	09YB53
12	109.5	4.6	824	2	09F216
13	109	4.6	821	2	059241
14	106.5	4.5	602	11	063485
15	106	4.5	798	4	09UN32
16	102.5	4.5	798	4	09UBK2
17	102.5	4.3	1686	4	000443
18	102.5	4.3	1686	4	000443
19	99.5	4.2	1509	11	061194

20	98.5	4.2	467	4	09UM5	09um5 homo sapien
21	98.5	4.2	467	4	09BRY2	09bry2 homo sapien
22	98.5	4.2	576	2	09L115	09l115 streptomyc
23	98.5	4.2	788	10	09SCV5	09scv5 arabidopsis
24	98.5	4.2	1658	11	061182	061182 mus musculu
25	98	4.2	435	2	09X9C0	09x9c0 streptococc
26	98	4.2	871	4	09Y5C3	09y5c3 homo sapien
27	98	4.2	938	4	09Y5F7	09y5f7 homo sapien
28	97.5	4.1	797	2	09RDM6	09rdm6 lactobacill
29	97	4.1	761	12	09WT88	09wt88 lt virus. o
30	96	4.1	422	2	09RC26	09rc26 streptomyc
31	95.5	4.0	389	2	074515	074515 synechocyst
32	95	4.0	523	4	09H856	09h856 homo sapien
33	95	4.0	739	5	09Y9E6	09y9e6 drosophila
34	94.5	4.0	476	5	09B160	09b160 caenorhabd1
35	94.5	4.0	530	5	045879	045879 caenorhabd1
36	94.5	4.0	683	2	09A6U7	09a6u7 caulobacter
37	94.5	4.0	749	2	059154	059154 anaerocellu
38	94.5	4.0	772	2	09Z3V2	09z3v2 pseudomonas
39	94.5	4.0	1046	2	084941	084941 streptococc
40	94.5	4.0	1641	2	09PDX7	09pdx7 xylella fas
41	94	4.0	454	11	09E054	09eq54 mus musculu
42	94	4.0	455	4	012875	012875 homo sapien
43	94	4.0	765	2	054183	054183 streptomyc
44	93.5	4.0	564	3	094727	094727 schizosacch
45	93.5	4.0	664	3	09C122	09c122 piromyces s

ALIGNMENTS

RESULT	ID	Query Match	Score	DB 2:	Length	Matches	Conservative	Pred.	No. 6:	Mismatches	Indels	Gaps
09K0U7	09K0U7	PRELIMINARY;	PRT;	382 AA.								
AC	09K0U7	09K0U7										
DT	01-OCT-2000 (TREMBLrel. 15, Created)											
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)											
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)											
DE	PHOSPHOLIPASE A1, PUTATIVE.											
GN	NMB0464.											
OS	Neisseria meningitidis (serogroup B).											
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.											
OX	NCBI_TaxID:491.											
RP	SEQUENCE FROM N.A.											
RC	STRAIN-MC58 / SEROGROUP B;											
RC	MEDLINE-20175755; PubMed-10710307;											
RA	Pettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,											
RA	Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,											
RA	Nelson W.C., Gynn M.L., Debey R., Peterson J.D., Hickey E.K.,											
RA	Hart D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,											
RA	Mason T., Ciecko A., Parksey D.S., Blair E., Ciltone H., Clark E.B.,											
RA	Cotton M.D., Uterback T.R., Khouri H., Qin H., Yamathavan J.,											
RA	Gill J., Scariato V., Maignani V., Piazza M., Grandi G., Sun L.,											
RA	Smith H.O., Fraser C.M., Moxon E.R., Rappoli R., Venter J.C.,											
RT	"Complete genome sequence of Neisseria meningitidis serogroup B strain											
RT	MC58."											
RL	Science 287:1809-1815(2000).											
DR	EMBL: AE002403; AAF40901.1; -.											
DR	TIQR: NMB0464.											
DR	InterPro: IPR003187; PLAI.											
DR	Pfam: PF02253; PLAI: 1.											
KW	Complete proteome.											
SQ	SEQUENCE 382 AA; 42714 MW; B46BA802F062E836 CRC64;											
QY	87	INCSALNDMDIACVDTLVHGETPAVI-----KTKRSIRLDERTIMQTI-KGKPOVYQE 140	34.5%	Score 815;	DB 2:	Length 382;						
DB	33	LCQCALNDVNTLACVTRIFRAQLPSSAGGEGESKAVLWTEVRSSLDKGEAVIYVEK 92	44.7%	Pred. No. 6:	5e-63;	Mismatches 111;	Indels 24;	Gaps 8;				

Qy	141	TTDPIELMGNKGMJLTKDKAQLEIYAKKQFPLPLSDDDDRNN-TPLMSSRPNNMYPL	199
Db	93	GGDAL-----PAASAGETADITPLSLMDLKDNDKGLGLGVREHNNPYLML	139
Qy	200	IFMHCKPMSRNPSTPSHEAR-OFTNEFRAPLEKFOVSVKKAABEDLWGTSDMLMFGYTOO	258
Db	140	LMYNNSPMYAGSPTPGTYTVEDEKGGQOKRAETKLQVSEFKSLIADDELKTKTADLMFGYTOR	199
Qy	259	SHMOJFN-GKNSRPFRRVNDYQPELFTLQPYVSDLPWGCKYRMIGMGAVHHNSSESAKLSR	317
Db	200	SDMOJYNGRKSAPFRNNDYKRELFELTQPYKADLPFGGRLRLMAGRVHNSQNSQRESR	259
Qy	318	SMNRAYLMAHEMKNLVYMPRIWGRIFKEEGSGSQPDNDPILDYEGYGDVRFPLYOLENKS	377
Db	260	SMNRIFYAAGAEKMLTIVIPRWVRARDQ--SGDK--NDNPIDADYWGSDVKLQYRLINDRQ	317
Qy	378	NISGTVKRNPPSSGCGALDLDVYVPLGKGISGYFOIDFGQSLIDYVHEATSTFSGVGLMLN	437
Db	318	NVYSYLRNPKPTGCAIEAAATPPIKCKLGVVAGFGHGSGLIDYVNHKNGIGIGIGLMEFN	377
Qy	438	DMWGL 442	
Db	378	DLDDGI 362	

Query Match	34.5%	Score 815	DB 2	Length 409
Best Local Similarity	44.7%	Pred. No. 7.1e-63		
Matches 163	Conservative 67	Mismatches 111	Indels 24	Gaps
QY 87	INCSALNDIMLACVYDPLVHGETPAVI-----KTKRSTRIDETIMQTI-KGKPOVYIOE	140		
DB 60	LOCALTDVNTVTLACDRFFAALQBPSSAQEQESKAVLNTLETYVASSLIDKGCAGAVTVKR	119		
QY 141	TTDPITFLMGNGEKMLTKKDAKOLEYAAKQFTPLSLISFDDLRRN-TPLWSSRPHPNMYVLP	199		
DB 120	GGDAL-----PADSAGETADIVTPLSLMYDLDKNDLRELLLVREHNPMYVLP	166		
QY 200	IFMHGKPRSPPTPSHEAR-QETPNEFRAPBLKFOYSAVYAKAEDLDMGDSDLMEGYTOO	258		
DB 167	LMYNNSPNAPSPSPRGTTVOEKFGQOKKAETKLYQSFSEKSLAEDLFTKRRADLMFGYTOR	226		

[illegible]

RESULT	3
ID	09CL22
AC	PRELIMINARY;
PR	306 AA.
DT	01-JUN-2001 (TREMBlrel. 17, Created)
DR	01-OCT-2000 (TREMBlrel. 17, Last sequence update)
DE	01-JUN-2001 (TREMBlrel. 17, Last annotation update)
GN	HYPOTHETICAL PROTEIN PM1426.
OS	Pasteurella multocida.
OC	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OX	Pasteurella.
RN	NCBI_TaxID=747;
RP	[1]
RC	SEQUENCE FROM N.A.
RA	STRAIN=PM70;
RL	MEDLINE=21145866; PubMed=11248100;
RT	May B.J., Zhang Q., Li L.L., Paustian M.L., Whitlam T.S., Kapur V.;
SR	"Complete genomic sequence of Pasteurella multocida pm70.";
DR	Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
KW	EMBL; AE006179; AAK03510.1; .
SQ	InterPro; IPR003187; PLAI1.
	Pfam; PF02253; PLAI1.
	Hypothetical protein; Complete proteome.
	SEQUENCE 306 AA; 35580 MW; EAF3DE8C1C22B26E CRC64;
Query Match	15.8%; Score 373.5; DB 2; Length 306;
Best Local Similarity	37.6%; Fred. No. 1.6e-24;
Matches 86; Conservative 43; Mismatches 91; Indels 9; Gaps	
OY	214 SHEAQFTFPNE-FRAPHKLFQVSVVKAAEDLMGTDSDLMFGYTQOSHQIFNGKNSRF 272
DB	81 TYSNNHFLKRETYQDDEIKFKISALPLMRGLGNNSVLAAYTKSWQLSNVDSSPF 140
OY	273 RVHYDQPEFIETLPQVSYDSLPMWDGKVMMIGMGAHVHSNG--ESAKLSRSNRYAYLMAGMEW 330
DB	141 RETNEPQQELFLMKTKQVSLPFCMTLDQVETGINHQSNCRGDAEKLRSRNRYLYBASAIK 200
OY	331 KNLUTMPRIKWCIEFKEGSSOPDDNPDLIDYYGYGDVR-LQLEKKSISGTVAIRPS 389
DB	201 QMWTEIKPWWIPIPK---AKNDNDPDITKYRGHFDAVALGYVYHDHQFLSG-HYNPIS 255
OY	390 GCAGVALDYVVYPLKGISGYSFOIPGGYGOSLLIDYHNHEATSFVGMLIND 438
DB	256 NKGGLEASTSYPTKNIRFYTYNGYGESLLIDYOQRRIORIGIGISLNN 304
RESULT	4
ID	09PMU8
AC	PRELIMINARY;
PR	329 AA.
DT	01-OCT-2000 (TREMBlrel. 15, Created)
DR	01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE	01-JUN-2001 (TREMBlrel. 17, Last annotation update)

DE PHOSPHOLIPASE A (EC 3.1.1.32).
GN PLDA OR CUI351.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OX Campylobacter.
NCBI_Taxid=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NCTC 11168;
RA MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Kelley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Felwell T., Holtroyd S.,
RA Jørgensen K., Kariyasekera A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rance M.A., Rutherford K.M., Van Vleet A.H.M.,
RA Whitehead S., Barrell B.G.,
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences."
RL Nature 403:665-668(2000).
DR EMBL: AL139078; CAB3778.1; -
DR InterPro: IPR003187; Pfam:
DR Pfam: PF02253; Pfam: 1
DR PRINTS: PRO1486; PHPLIPASEA1.
KW Complete proteome.
SQ SEQUENCE 329 AA; 38880 MW; A32731F2B751AC44 CRC64;

Query Match 15.5%; Score 366; DB 2; Length 329;
Best Local Similarity 33.2%; Pred. No. 8.2e-24;
Matches 105; Conservative 48; Mismatches 125; Indels 38; Gaps 13;

OY 130 IKGAPVYVETTPDPIFLMGNEKGLTKKDAKOLEYAKOFTPLSDIDBRNNT--PLM 187
DB 43 LKNSSVLSQDQNNSSQATOTQNTTKEKQ-----DPSRLALAVYLGENSEFNDLG 96
OY 188 SSRPNPVPYVPIFMGKPNRSPPTPSHEARQTPNEFRABELKQVSVKAAEDLMGT 247
DB 97 IS-SYKMYFLP-FAYSPNSLGVNNKSEA-----KQLSVKRLKFLFNILGL 141
OY 248 DSDLMFGTQOOSHQIENGKSRPRVHDYQPEIFLTPYV-SDLPMGKRYMIGMAVH 306
DB 142 DEKYIATQTSWQIY--EHSSPRETNYQPEFIDLPYLKDYEFNNLR--VGILH 196
OY 307 HSNSESAR--LSRSNRAYILMAGMEKMLTYVPRIMGIFEGSGSDPDNDILDYGY 364
DB 197 ESNKGDENLOSRSNRYSTALIKYKFLVPRIMYRI---PEKKKDDNDPAILHYGN 253
OY 365 GDVAFELYOLEKSNISGTVRYNPR--SGKALQIDYVPL-GKGISGYFOIGYQSLI 421
DB 254 EDVNLAY-LGDYFINLMLRNKLFHNKKAIDYDLYGDIENNIGIYWLQYFNGYGESLI 312
OY 422 DYNHEATSFYGLMLN 437
DB 313 DYNKHLQRLSTGFLIS 328

RESULT 5
OQSID7
ID OQSID7 PRELIMINARY; PRT; 292 AA.
AC OQSID7;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE PHOSPHOLIPASE A.
GN PLDA.
OS Yersinia pseudotuberculosis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX Yersinia.
NCBI_Taxid=633;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-YPIII P1B1;
RA Kariyasekera A.V., Winzeler E.A., Williams K.J., Oyston P.C.,
RA Tibball R.W., Wren B.W.;

RT "Biochip-based signature-tagged mutagenesis: identification and
RT characterisation of Y. pseudotuberculosis gene plda essential for
RT virulence in mice."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ245393; CAB51586.1; -
DR InterPro: IPR003187; Pfam:
DR Pfam: PF02253; Pfam: 1
DR PRINTS: PRO1486; PHPLIPASEA1.
SQ SEQUENCE 292 AA; 33758 MW; 8E712D908ACB6BA5 CRC64;

Query Match 15.3%; Score 360.5; DB 2; Length 292;
Best Local Similarity 39.7%; Pred. No. 2.1e-23;
Matches 83; Conservative 34; Mismatches 86; Indels 9; Gaps 5;

OY 229 EIKFQSVKVKAAEDLMGTDSDLMFGYTOOSHQIENGKSRPRVHDYQPEIFLTPYV- 287
DB 86 EVKFLSLAFPIFMGKPNRSPPTPSHEARQTPNEFRABELKQVSVKAAEDLMGT 145
OY 288 YSDLPMDGKVRMIGMAVHSHNGESAKLSRSNRAYILMAGMEKMLTYVPRIMGRIKFG 347
DB 146 YELAGW--TFREVERGFPHQNSGKADPTSRSNRYTRYMAORGMLDLPKFWRIPESD 203
OY 348 SGSDPDNDILDYGYGVDFRFLYQV-ENKSNISGTVRYNPRSGKALQIDYVPLGKI 406
DB 204 S---KDDNPDIKKYMGYRLKAGYALGDSVFSLDG--RNNWTVGAGAEKMGSTYITKHV 258
OY 407 SGYFOIFQYQSLIDYDYNHEATSFYGLMLN 438
DB 259 RFTQVFSYSGESMIDYNTROTAVGVGIMLD 290

RESULT 6
O9L6N9
ID O9L6N9 PRELIMINARY; PRT; 289 AA.
AC O9L6N9;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE PLDA PROTEIN.
GN PLDA.
OS Salmonella typhimurium LT2.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX Salmonella.
NCBI_Taxid=99287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SGSCI412;
RA Maabu;
RT "The Salmonella typhimurium Genome Sequencing Project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-SGSCI412;
RA Waterston R.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF233324; AAF33435.1; -
DR InterPro: IPR003187; Pfam:
DR Pfam: PF02253; Pfam: 1
SQ SEQUENCE 289 AA; 32967 MW; DA97F5E1651C49C6 CRC64;

Query Match 15.3%; Score 360; DB 2; Length 289;
Best Local Similarity 39.7%; Pred. No. 2.3e-23;
Matches 85; Conservative 29; Mismatches 90; Indels 10; Gaps 4;

OY 226 RABELKFOYSVKVKAAEDLMGTDSDLMFGYTOOSHQIENGKSRPRVHDYQPEIFLQ 285
DB 83 RKDEVKFOULAFPIFMGKPNRSPPTPSHEARQTPNEFRABELKQVSVKAAEDLMGT 142
OY 286 PV-YSDLPMDGKVRMIGMAVHSHNGESAKLSRSNRAYILMAGMEKMLTYVPRIMGRIF 344
DB 143 ATDYRFAGW--TLRDVEKGNHDSNGRSDPTSRSNRLTRYLMAGNGLVYKRWYI- 199

```

QY 345 KEGSGSQPDNDPILDYGGVDFLYQLEKNKNSIGTVRYNRSRGALQDLYVYPLGK 404
    |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB 200 -----GSTPDNDPITFYKMGYQKIGYHL-GEAVLSAKQGYNMNTGGAEGVLSYVTK 253
QY 405 GISGYFOIFOGYGSGLIDYHNEATSFVGLMND 438
    ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB 254 HVRLYTVYSGESGLIDYHNEATSFVGLMND 287

RESULT 7
O924N8 PRELIMINARY; PRT; 292 AA.
AC O924N8;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
DE OUTER MEMBRANE PHOSPHOLIPASE A PRECURSOR (EC 3.1.1.32).
GN PLDA.
OS Enterobacter agglomerans.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Pantoea.
OX NCBI_TaxID=549;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94131966; PubMed=8300539;
RA Brok R.G., Brinkman E., van Boxtel R., Bekkers A.C., Verheij H.M.,
RA Tomassen J.;
RT "Molecular characterization of enterobacterial plda genes encoding
RT outer membrane phospholipase A.;"
RN J. Bacteriol. 176:861-870(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120145; PubMed=9921577;
RA Brok R.G., Boots A.P., Dekker N., Verheij H.M., Tomassen J.;
RT "Sequence comparison of outer membrane phospholipases A: implications
RT for structure and for the catalytic mechanism.;"
RL Res. Microbiol. 149:703-710(1996).
DR EMBL: AF034414; AAD03498.1;
DR InterPro: IPR003187; PLA1.
DR Pfam: PF02253; PLA1; 1.
DR PRINTS: PR01486; PHPLIPASEA1.
KW Signal; Hydrolase.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 292 OUTER MEMBRANE PHOSPHOLIPASE A.
SQ SEQUENCE 292 AA; 33719 MW; B75516D093B2BEEA CRC64;

Query Match 15.1%; Score 356.5; DB 2; Length 292;
Best Local Similarity 40.1%; Pred. No. 4.7e-23;
Matches 85; Conservative 31; Mismatches 87; Indels 9; Gaps 5;

```

```

DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
DE PHOSPHOLIPASE A.
GN PLDA.
OS Campylobacter coli.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=195;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=DA585;
RX MEDLINE=97230284; PubMed=9119448;
RA Grant K.A., Belandria I., Dekker N., Richardson P.T., Park S.F.;
RT "Molecular characterization of plda, the structural gene for a
RT phospholipase A from Campylobacter coli, and its contribution to cell-
RT associated hemolysis.;"
RL Infect. Immun. 65:1172-1180(1997).
DR EMBL: Y11031; CA471915.1;
DR InterPro: IPR003187; PLA1.
DR Pfam: PF02253; PLA1; 1.
DR PRINTS: PR01486; PHPLIPASEA1.
SQ SEQUENCE 297 AA; 34998 MW; 04B54A7BCA3764CE CRC64;

Query Match 14.5%; Score 342.5; DB 2; Length 297;
Best Local Similarity 32.4%; Pred. No. 8e-22;
Matches 97; Conservative 37; Mismatches 108; Indels 57; Gaps 11;

QY 178 DLDRNTPLMSSRPNDYVLPFMCKPNRSPT-----P 213
    ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB 16 DLKENNASLSRIKHETQ-----NTKPTSTKEDFSRIALANYGENSSENPUGIS 66

QY 214 SHEARQTP-----NEFRAPELKEQYSVKKAAEDLMDGTSDLMFGYTOOSHQIF 264
    ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB 67 SYKNNTFLPRAYSGSLGKNGKRTKMFQUSIKRFLFEDLGLGEKYVGTOTSMQ-- 124

QY 265 NGRNSRPFRVHDYQPELEFLOPV-YSDLPWDGKVRMIGMAVHNSGESAK--LSRSWR 321
    ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB 125 NYKHSSPFRFNQPEFEVDICLHFEDYKFLNLR--VGILHSNKGDEPNESRSMNR 181

QY 322 AYLMAKMEKNLYMPFRMGKIRFREGSGSQPDNDPILDYGGVDFLYQLEKNSISG 381
    ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB 182 IVASSVFLYORFLEVPRIWRI--PENSDDDNPELTHMGWEDIN-IGSLGNDYFINL 237

QY 382 TVRYNP--RSGKGLQDLYVPL-GKISGYFOIFOGYGSGLIDYHNEATSFVGLMND 437
    ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB 238 MLRNNDLFDHDKGAVQVDIGDIFDNGIYIWLQTFNGYGSGLIDYKRLRLSTAFLLIS 296

RESULT 9
O25241 PRELIMINARY; PRT; 355 AA.
AC O25241;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
DE PHOSPHOLIPASE A1 PRECURSOR (DR-PHOSPHOLIPASE A).
GN HP0499.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=26595 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,

```


RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Mallin E.,
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
 RA Venter J.C.;
 RT "The complete genome sequence of the gastric pathogen *Helicobacter*
 RT *pylori*.";
 RL Nature 388:539-547(1997).
 DR EMBL: AE000564; AAD07564.1; -
 DR TIGR: HP0499; -
 DR InterPro: IPR003187; Pfam: 1.
 DR Pfam: PF02253; Pfam: 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 355 AA; 42486 MW; 461960F15E65AB0A CRC64;

Query Match 10.4%; Score 246.5; DB 2; Length 355;
 Best Local Similarity 25.1%; Pred. No. 2.5e-13;
 Matches 82; Conservative 41; Mismatches 117; Indels 87; Gaps 11;

QY 157 KKDAKOLEYAKOFTPLSLSFDDLRNNTPLMSSRPHNPMVYLPIFMHGKPRKSPNTPSHE 216
 DB 69 KKYLMMDYLGTYFLPYHSHF-----TPIFGWYHPNINP----- 102
 QY 217 AROFTPNEFRAPELKFOVSVKKAEDLWGTDSDLMFGYTOOSHMOJFNKNSRPFVYHD 276
 DB 103 ---YORNEF-----KFOISFRVPRVFRHILMTKGLTYLATYOTMFOIYNDPQSAFMRMN 154
 QY 277 YQPEIFLQPYVSDLPMDGK---RMIGCAVHNSG-ESAKLSRSMNRAYLMAEMWKN 332
 DB 155 FMPFLIYVYPI-NFKPFGKIGNESEIWMQHSNGVGAQCYQPFNK---EGNPNQ 209
 QY 333 LTWVPRF-----WGRIFKESGSGSP-----DONPD 357
 DB 210 FPGQPVIVKQDNGKQDVWGGCRSVSAGQRPVFLWMEKGLKIMVAYWPVYDQSPN 269
 QY 358 ILDYGYGADVRELY-----OLENKSINSGTVYRNPNSGKALQLDVYVPLKGISGYF 410
 DB 270 LIDWYGKGNKIDYRGRHHELDYDIFQYWRD--RMHGARLGLTYRINPFVGYA 327
 QY 411 QIFQGYGOSLIDYNHEATSFVGLMLN 437
 DB 328 QWENGXGDLGYEYDVFNSRIGVGIRLN 354

RESULT 10
 Q9ZLX5 PRELIMINARY; PRT; 355 AA.
 AC Q9ZLX5;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE PUTATIVE PHOSPHOLIPASE A1.
 GN PLDA OR JHP0451.
 OS *Helicobacter pylori* J99 (Campylobacter *pylori* J99).
 OC Bacteria; Proteobacteria; epsilon subdivision; *Helicobacter* group;
 OC *Helicobacter*.
 NC NCBI_Taxid=85963;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99120557; PubMed=9923682;
 RA Alm R.A., Ling L.-S.L., Moll D.T., King B.L., Brown E.D., Dolg P.C.,
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
 RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human
 RT gastric pathogen *Helicobacter pylori*.";
 RL Nature 397:176-180(1999).
 DR EMBL: AE001479; AAD06029.1; -
 DR InterPro: IPR003187; Pfam: 1.
 DR Pfam: PF02253; Pfam: 1.
 KW Complete proteome.
 SQ SEQUENCE 355 AA; 42329 MW; B3CE9810EBA2FDC9 CRC64;

Query Match 10.2%; Score 240.5; DB 2; Length 355;
 Best Local Similarity 25.1%; Pred. No. 8.3e-13;
 Matches 83; Conservative 44; Mismatches 109; Indels 95; Gaps 13;

QY 157 KKDAKOLEYAKOFTPLSLSFDDLRNNTPLMSSRPHNPMVYLPIFMHGKPRKSPNTPSHE 216
 DB 69 KKYLMMDYLGTYFLPYHSHF-----TPIFGWYHPNINP----- 102
 QY 217 AROFTPNEFRAPELKFOVSVKKAEDLWGTDSDLMFGYTOOSHMOJFNKNSRPFVYHD 276
 DB 103 ---YORNEF-----KFOISFRVPRVFRHILMTKGLTYLATYOTMFOIYNDPQSAFMRMN 154
 QY 277 YQPEIFLQPYVSDLPMDGK---RMIGCAVHNSG-ESAKLSRSMNR----- 321
 DB 155 FMPFLIYVYPI-NFKPFGKIGNESEIWMQHSNGVGAQCYQPFNK---EGNPNQ 213
 QY 322 -----AYLMAG-----MENK--NTWVPRFGRIFKESGSGSPDD- 354
 DB 214 PVYKVDNGKQDVWGGCRSVSAGNALCFVLWMEKGLKIMVAYWPV-----PYDQ 265
 QY 355 -NPDIDYGYGADVRELY-----OLENKSINSGTVYRNPNSGKALQLDVYVPLKGI 406
 DB 266 SNPQIDWYGKGNKIDYRGRHHELDYDIFQYWRD--RMHGARLGLTYRINPFV 323
 QY 407 SGYFOIFQGYGOSLIDYNHEATSFVGLMLN 437
 DB 324 GYVQWENGXGDLGYEYDVFNSRIGVGIRLN 354

RESULT 11
 Q9XB53 PRELIMINARY; PRT; 278 AA.
 AC Q9XB53;
 DT 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
 DE KDOI.
 GN KDOI.
 OS *Erwinia carotovora*.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Pectobacterium.
 NC NCBI_Taxid=554;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 39048; GS101;
 RX MEDLINE=98065591; PubMed=9402024;
 RA McGowan S.J., Seabright M., O'Leary S., Hardie K.R., Williams P.,
 RA Stewart G.S., Bycroft B.W., Salmond G.P.;
 RT "Analysis of the carbenpenem gene cluster of *Erwinia carotovora*:
 RT definition of the antibiotic biosynthetic genes and evidence for a
 RT novel beta-lactam resistance mechanism.";
 RL Mol. Microbiol. 26:545-556(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 39048; GS101;
 RX MEDLINE=98276484; PubMed=9614345;
 RA McGowan S.J., Bycroft B.W., Salmond G.P.;
 RT "Bacterial production of carbenpenems and clavams: evolution of beta-
 RT lactam antibiotic pathways.";
 RL Trends Microbiol. 6:203-208(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 39048; GS101;
 RA McGowan S.J.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U17224; AAD38237.1; -
 SQ SEQUENCE 278 AA; 31341 MW; 63769F4F3550E1B7 CRC64;

Query Match 5.1%; Score 121; DB 2; Length 278;
 Best Local Similarity 22.0%; Pred. No. 0.016;
 Matches 71; Conservative 49; Mismatches 125; Indels 78; Gaps 15;

```

QY 70 EHEBELTTLAENTNTMINSALNODIMRLACVDYLHGSETPAVITIKRIRLDETWO- 128
Db 10 EHATLOTTELKRFKLEJLEFTFPOQYMTYSHIDRIYVG--GIMPVDSITFDDGIGK 66
QY 129 -----TKGRQVUYQETDTPDIFLMGNEKMLTKKADOLEYAKOETPL 173
Db 67 FGWYFELERRELGLINGGPAKVIIDGTS--EYVGHEEALYVGKGAKALAFS----- 116
QY 174 SLSEFDLRNNTPLWSSRPNHPMYLPIFGMGKPNRSPNTPSHEAROFTRPEFAPELKO 233
Db 117 -SLDSAKPAKLYUNSAIPAHHVPEFTRLITQDAIKAPLDGVKCNCKRTICKYLVPFE- 171
QY 234 VSVAVKAAEDMLGDSOLMGYGT---QOSMOLFNGKNSRPFVNDYQPIFLTOPIYSD 290
Db 172 ----VEFCQ-----LSMGLTRLAGSNW-----NSMPYTHHERMVEYF---YED 209
QY 291 LPMWDGKVRMIGMGAVNHS-----NGESAKLSRSNRNAYILMAGEMKNLTYMPRINGRIFK 345
Db 210 MAEDTILFIHM-MGEPEHTRHLYMHNEOAVISPMS--IHTGVTKRYAE--IWMGI-- 260
QY 346 EGSQSGOPDDNPDLIDYGYGQVVR 368
Db 261 -GENUTFDD---MDHTAMDLR 278

```

RESULT	12			
09F216				
ID	09F216	PRELIMINARY:	PRT:	824 AA.
AC	09F216;			
DT	01-MAR-2001 (TREMBLrel. 16, Created)			
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	CELLULOSE.			
OS	Bacillus sp.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;			
OC	Bacillus/Staphylococcus group; Bacillus.			
OX	NCBI_TaxID=1409;			
NP	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-KSM-S237;			
RA	MEDLINE-21036886; PubMed-11193393;			
RA	Hakemada Y., Hatada Y., Koike K., Yoshimatsu T., Kawai S.,			
RA	Kobayashi T., Ito S.;			
RT	"deduced amino acid sequence and possible catalytic residues of a			
RT	thermostable, alkaline cellulase from an alkaliphilic Bacillus			
RT	strain.";			
RL	Biosci. Biotechnol. Biochem. 64:2281-2289(2000).			
DR	EMBL; AB018420; BAB9360.1;			
DR	InterPro: IPR001547; Glyco_hydro_F5.			
DR	Pfam: PF00150; cellulase: 1.			
DR	PROSITE: PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN_1.			
SO	SEQUENCE 824 AA; 91364 MW; 65FA840FE1D729B9 CRC64;			

Query Match	4.68	Score 109.5	DB 2	Length 824
Best Local Similarity	19.78	Pred. No. 0.76		
Matches 91	Conservative 57	Mismatches 146	Indels 167	Gaps 27

```

QY      30 PVAIVDEVRKSN-----LGQDNELLGVGASQASASTIDRANPL-DEHEPELY-----TT 78
DQ      1  - - - - -
DQ      25 PALLAAREGNTREDNEFKHLGNDNVKRPSEAGALDLDGVDSQOMTLVDQHGSEIKLRGMSH 84
QY      79 ALEKKTALINCSAL-----NODIMFLACYDTLVHGE-----TPRAVIATKK-----RST 120
DQ      85 GLQFPEPFLINDNNAVKALSNDSWDSMIRLAMY-----VEENGATNPBELIKORIVIDIELAI 140
QY      121 RLDEETI---WQT-----IKGRPOVYVOYQETTOR-----IF 146
DQ      141 ENDMYIVLYDMVHVAHAPGRDPROVYAGARDFREIETALYPPNNPHITTELANEPSSNNNGAG 200
QY      147 LMGNEKGMLTKKDAKOLEVAAKQETPLSLSFIDLDRN-----NTPLWSSRP-----H 192

```

```

Dh 201 IPNNEEGKRAVK-----EADPIVEMRKSGNADNDIIIGVSPNWSQRPDLADNPIDDH 255
Qy 193 NPMYVPLEFMHGKRNKRSNPNTSHAROPTNREFAPBLK-----FOVSYKVAABEDMCT 247
Dh 256 HPMYTVVHEHYGSHAASIESYPSF-----TPJSEGNVMSNTRYALENGAVAFATE--WGT 308
Qy 248 -----DSDLWFGYTOQ-----SHWOIFNGKNS-----RPFV----- 274
Dh 309 SQASGDGPFYDEADVWIEFLINENIMISMANWSLTN-KNEVSGAFTPEELCKSNATNLDDG 367
Qy 275 --HDQP-ELFLT-----OPV-----YSDLPW--DGKVRMTMGAVHHSNCE 311
Dh 368 PDHWAPPEELSLGSEYVHARIKYVNEPIRDTKTKLWLMDFNGTQO--GFGVNSDSPNK 425
Qy 312 SAKTSRSNRRAYLAGME-----WKNLTPWPRIMGR 342
Dh 426 ELIADVDENNNTLTKVSGLDVSNDSVDGFMFMANARLSANGMCK 466

```

RESULT	13			
059241				
ID	059241	PRELIMINARY;	PRT;	821 AA.
AC	059241;			
DT	01-NOV-1996 (TEMBLrel. 01, Created)			
DT	01-NOV-1996 (TEMBLrel. 01, Last sequence update)			
DT	01-JUN-2001 (TEMBLrel. 01, Last annotation update)			
DE	ENMO-1,4-BETA-GLUCANASE (EC 3.2.1.4) (CELLULASE) (ENDOGALACTANASE),			
DE	(CARBOXYMETHYL CELLULASE).			
OS	Bacillus sp.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;			
OC	Bacillus/Staphylococcus group; Bacillus.			
OX	NCBI_Taxid:1409;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-KSM-64;			
RA	Sumitomo N., Ozaki K., Ito S.;			
RL	submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.			
CC	-1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC			
CC	LINKAGES IN CELLULOSE.			
CC	EMBL; M64963; AAA73189.1; -.			
DR	HSSP; 085465; 1A3H.			
DR	InterPro; IPR001547; Glyco_hydro_F5.			
DR	Pfam; PF00150; cellulase; 1.			
DR	PROSITE; PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN1.			
KW	Hydrolase; Glycosidase			
SQ	SEQUENCE 821 AA; 90910 MW; 73d438fEFD40b5C CRC64;			

Query Match	4.6%;	Score 109;	DB 2;	Length 821;
Best Local Similarity	19.7%;	Pred. No. 0.83;		
Matches	91;	Conservative	51;	Mismatches 150;
			Indels	170;
			Gaps	24;

QY	30	PVA	F	E	D	E	V	R	K	S	K	N	D	---	I	G	O	N	E	L	L	I	G	V	S	A	N	O	S	A	S	T	A	N	P	---	D	E	H	E	B	E	L	---	T	T	78							
Db	24	P	T	A	A	E	G	T	R	E	D	N	P	---	K	H	L	G	D	N	V	K	R	P	S	E	A	L	I	D	O	E	V	D	G	M	T	L	V	D	O	H	G	E	K	I	D	R	G	M	S	T	H	83
QY	79	A	L	E	K	T	M	I	N	S	A	L	---	N	O	I	M	R	A	C	---	D	L	V	I	G	E	T	R	A	V	A	K	T	R	S	I	R	D	E	T	---	126											
Db	84	G	L	O	M	P	E	L	L	N	A	V	K	---	K	A	L	A	N	D	E	S	M	I	R	L	A	M	V	G	E	N	G	A	S	N	E	L	L	K	I	S	V	I	G	I	D	A	T	E	N	D	143	
QY	127	---	W	O	T	---	Y	O	E	T	T	D	---	I	K	R	O	V	U	Y	O	E	T	T	D	---	I	F	L	M	G	N	150																					
Db	144	Y	V	I	D	M	H	N	A	R	C	D	R	P	---	Y	A	G	A	E	D	F	F	R	D	I	A	L	Y	N	N	H	I	I	E	L	A	N	E	S	S	N	N	G	A	G	I	P	N	203				
QY	151	E	K	M	L	T	K	D	A	K	O	L	E	Y	---	A	K	O	T	T	P	L	S	F	D	L	R	---	N	T	P	L	M	S	S	R	---	196																
Db	204	E	E	G	M	N	A	V	---	E	X	A	R	P	---	I	E	M	L	D	S	G	N	A	D	N	I	I	I	V	E	S	P	M	S	O	R	P	D	L	A	N	P	I	D	I	H	T	M	258				
QY	197	V	L	P	T	---	H	G	K	P	N	S	---	P	N	P	S	H	E	A	R	G	T	P	E	R	F	A	R	L	K	O	V	S	K	V	A	A	D	I	G	---	247											
Db	259	T	V	H	I	T	G	S	H	A	S	T	E	S	---	P	R	E	T	P	E	T	N	S	R	G	N	M	S	T	R	I	A	---	L	E	N	G	A	V	A	F	A	T	---	M	G	T	S	O	A	N	G	312
QY	248	---	D	S	D	L	M	F	G	T	O	---	S	H	M	O	I	F	N	G	K	N	---	R	P	F	R	V	---	275																								

Db 313 DCGPYFDEADVIEFLMENNISMWNSLTN-KNEVSAFFPELFGKSNATSLDGPDPYV 371
Qy 276 -----DYQPEIFLTPYSDLPW---DGKVMIGMGAHHNSGES- 312
Db 372 VPEELSLSGEYVARIKGVNEP---IDRTKYFLVMDPFGNGFGV-----NGDSP 422
Qy 313 -----AKLSRWNRAYLMAGMEKNLTPMPCIMR 342
Db 423 VEDVYIENAGALKSLGDSNDVSEGNMYMANRLSADGCK 464

RESULT 14

063485 PRELIMINARY; PRT: 602 AA.
AC 063485;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE C-RAF ACTIVATED ONCOGENE FUSION PROTEIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=67172791; PubMed=3550433;
RA Ishikawa F., Takaku F., Nagao M., Sugimura T.;
RT "Rat c-raf oncogene activation by a rearrangement that produces a
RT fused protein.";
RL Mol. Cell. Biol. 7:1226-1232(1987).
CC - SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: M15428; AAA42002.1; -;
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_kin_actsite.
DR Pfam: PF00069; pkinase.1.
DR SMART: SM00221; STRYK: 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_DOM; 1.
DR AMP-binding; Kinase; Oncogene; Serine/threonine-protein kinase;
KW Transferrase.
SQ SEQUENCE 602 AA; 69209 MW; 88B7BFA90FFB02AC CRC64;

Query Match 4.5%; Score 106.5; DB 11; Length 602;
Best Local Similarity 19.0%; Pred. No. 0.88;
Matches 92; Conservative 74; Mismatches 172; Indels 145; Gaps 24;

Qy 23 QAOAVPVPVAFVDEVRKNDLGONELLIGVQASSTDTANPLDHEPELYTTALEN 82
Db 81 ELEFLNNQVYVLETKNELTAQDRNL--GIQSOFTRAKEE---LEAEKRDILRT--N 131
Qy 83 KTMNLINCSALNODIMRLACVDTLVHGEFPAVYIKRSIRLDETIWQ--TIKGPQVYVOE 140
Db 132 ERLSQEVEYLETEDVKRL--NEKLESNT--TKGELDLKLDELQASDVYVYRERLEQF 186
Qy 141 -----TTPPIFLMGNEKG-----MLTKKDAKOLEYAAQFTPLS 174
Db 187 KELLHNQNSMLNTELTAKTDBELLALGRKQNEILELCTLENKKKEPDAIRSHSSASASA 246
Qy 175 LSFPLDNNNPPLMSSRRHPNRYVLPITMHGKPNRSPPTPSHEARQFTPNFRAPDLKFQV 234
Db 247 LS--SSPNLSPGWSQPKTP-----VPAQREPAFGSGTOEKNKIRPRGQDSSYYWEI 298
Qy 235 SVKKAEDLMGTDSDLMFGYTOOSHW-----QIFNGKNSRPVRVHDYOPEI----- 281
Db 299 EASEVMLSTRIGSGS---FGTVYKGAHGDVAVKILKVVDTPQDLQAFRNEVAVLKRTR 355
Qy 282 -----FLTQ-----PVYSDL--PMDGKVRMI-----GMGAVNH 307
Db 356 HVNILLFMGYTKDNLAIIVQWCGESGLYKHLHYQETKRFQWFLDIDARQAQGM DYLHA 415

Qy 308 SNGESAKLRSRWNRAYLMAGMEKN-----NLTPMPCIMRIFKEGSGSQ----- 351
Db 416 KNIILHRDKS--NNIFLEGLLVKIGDFGLATVSRW-----SGSQVQDPPIGSVLM 466
Qy 352 -----PDDNPDLI--DYGYGDVRYFLEENKSNISGTVARNRSGKALQLDYVPL 402
Db 467 APEVIRMQDNNPFSPQSDVYSYGIY--LYEL-----MTGELPYSHINRQDI-----IRMV 515
Qy 403 GKG 405
Db 516 GRG 518

RESULT 15

090N32 PRELIMINARY; PRT: 798 AA.
AC 090N32;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PPAR GAMMA COACTIVATOR-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Larrouy D., Vidal H., Andreelli F., Laville M., Langin D.;
RT "Cloning and mRNA tissue distribution of human PPARgamma coactivator-
RT 1.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF159714; AAD51615.1; -;
DR InterPro: IPR000504; RRM.
DR Pfam: PF00076; rrm; 1.
DR PROSITE: PS50102; RRM; 1.
DR SMART: SM00360; RRM; 1.
SQ SEQUENCE 798 AA; 91056 MW; F68F9768BD94E1F4 CRC64;

Query Match 4.5%; Score 106; DB 4; Length 798;
Best Local Similarity 24.1%; Pred. No. 1.5;
Matches 93; Conservative 36; Mismatches 115; Indels 142; Gaps 27;

Qy 3 VSLSTLTSLIPCFRILTAIQQAQAVPVPVAFVDEVRKNDLGONELLIGVQASST 62
Db 89 LAVLETLDLSLPV-----DEGLPSFDALTD-----GDVTTDNE-----ASPSSMP 129
Qy 63 D-TANPLDHEPELYTTAL--ENKTMILIN-CSALNODIMRLACVDTLVHGE-----T 110
Db 130 DGTPEPQAEERPSLKLKLLAPANTQLSYNECSGIS-----TONHANHNRIRTN 179
Qy 111 PAVIKTKRSIRLDETIWQ-----IKGKPQ-----VVOETPD-----PIFLMGNE 151
Db 180 PAIYKTEMS-----WNNKAKSICQOQKQRRPQSELSLKLTLTNDPPTKPFENRRSS 232
Qy 152 KGMLT--KKDAKOLEYAAQFTPLSFDLDRNNTPLMSSRRHPNM-----VYLP 199
Db 233 RDKCTSKKSHSQSOQHLOAKPTLSL-----PLTPESPNDPKSGFPENKTIERTLS 285
Qy 200 IFMNGKPNRS--PNTPSHEARQFTPNFERA--PELKFQV-----SVKKAEDLMGT--- 247
Db 286 VELSGTAGLTPPTTPPHKANO--DNPFASPKLSSCKTVVPPSKPKRYSESS--GTQGN 342
Qy 248 -----DSDLMNGYTOOSHWQIFNG-----KNSRP-----FRVHDY-----OPEIFL- 283
Db 343 NSTKKGPQSEL---YALSKSSSVLTGSGHEERKTRPRLRFGHDHYQGSINSTELLIN 399
Qy 284 -----TQPVYSDLPMDGKV 297
Db 400 ISOELQDROLEKNRDVSSD--WQGOI 423

Search completed: November 30, 2001, 14:26:27
Job time: 564 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:16:59 ; Search time 72.04 Seconds
(Without alignments) 454.475 Million cell updates/sec

Title: US-09-787-083-6

Perfect score: 2360
Sequence: 1 MKVSLSTLTLSCFALLA.....YHNEATSEGVGLMDMGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: A_Geneseq_1101.*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
23: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2360	100.0	442	AA1985270	BASB034 amino acid
2	2352	99.7	442	AA1985268	BASB034 amino acid
3	2340	99.5	442	AA1985271	BASB034 amino acid
4	2330	98.7	442	AA1985269	BASB034 amino acid
5	822	34.8	370	AA1985156	Neisseria gonorrhoe
6	815	34.5	370	AA1985157	Neisseria meningit
7	815	34.5	370	AA1985158	Neisseria meningit
8	815	34.5	374	AA1985159	Neisseria meningit
9	812	34.4	375	AA1985160	Neisseria meningit
10	246.5	10.4	355	AA1985161	H. pylori GHP0 172
11	239.5	10.1	356	AA1985162	H. pylori ORF 07ap

12	156.5	6.6	253	AA1985270
13	114	4.8	1685	AA1985271
14	109.5	4.6	800	AA1985272
15	109.5	4.6	822	AA1985273
16	108	4.6	824	AA1985274
17	107	4.5	157	AA1985275
18	104.5	4.4	537	AA1985276
19	104	4.4	798	AA1985277
20	104	4.4	798	AA1985278
21	104	4.4	798	AA1985279
22	100.5	4.3	682	AA1985280
23	99.5	4.2	1726	AA1985281
24	98.5	4.2	467	AA1985282
25	98.5	4.2	467	AA1985283
26	98.5	4.2	467	AA1985284
27	98.5	4.2	467	AA1985285
28	95.5	4.0	502	AA1985286
29	95.5	4.0	516	AA1985287
30	95	4.0	522	AA1985288
31	94	4.0	372	AA1985289
32	94	4.0	761	AA1985290
33	94	4.0	1115	AA1985291
34	93.5	4.0	467	AA1985292
35	92.5	3.9	1717	AA1985293
36	92	3.9	888	AA1985294
37	91.5	3.9	857	AA1985295
38	91	3.9	1227	AA1985296
39	90.5	3.8	1024	AA1985297
40	90	3.8	460	AA1985298
41	90	3.8	481	AA1985299
42	90	3.8	522	AA1985300
43	90	3.8	564	AA1985301
44	90	3.8	564	AA1985302
45	90	3.8	600	AA1985303

ALIGNMENTS

RESULT 1	
AA1985270	AA1985270 standard; Protein: 442 AA.
ID	AA1985270
AC	AA1985270;
XX	
DT	29-JUN-2000 (first entry)
XX	
DE	BASB034 amino acid sequence #3.
KW	Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KW	vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW	sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW	hearing loss; antibacterial drug.
XX	
OS	Moraxella catarrhalis.
XX	
PN	W0200015802-A1.
XX	
PD	23-MAR-2000.
XX	
PF	14-SEP-1999; 99WO-EP06781.
XX	
PR	14-SEP-1998; 98GB-0020002.
XX	
PA	(SMIK) SMITHKLINE BECHAM BIOLOGICALS.
XX	
PI	Ruelle J;
XX	
DR	WPI: 2000-271440/23.
XX	
DR	N-PSDB; AAA10702.
XX	
PT	Novel BASB034 polynucleotides and polypeptides from Moraxella
PT	catarrhalis used to prepare vaccines against bacterial infections

XX Claim 3; Page 68; 106pp; English.

CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC strain Mc2913. The invention relates to BASB034 polypeptides from
CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2959. The BASB034
CC polynucleotides and polypeptides may be employed as research reagents and
CC material for the discovery of treatments and diagnostics for diseases,
CC particularly human diseases. They are particularly used to diagnose and
CC treat M. catarrhalis infections. They can be particularly used for diagnosis of
CC disease, staging of disease, or determining response of an infectious
CC organism to drugs. The polynucleotides may be used as a source for
CC hybridization probes, and for screening of genetic mutations, serotype,
CC organism or strain identification, identification of mutations in BASB034
CC sequences, and as components of arrays which are useful for diagnostic
CC and prognostic purposes. The polypeptides can be used to produce
CC antibodies. The polypeptides can also be used in vaccine formulations,
CC and to identify agonists and antagonists. The polypeptides, antibodies,
CC agonists and antagonists (which are bacteriostatic) are used for the
CC treatment and prevention of diseases such as otitis media in infants and
CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
CC invasive diseases, and chronic otitis media with hearing loss. The
CC polypeptides, agonists and antagonists are also used for screening of
CC antibacterial drugs. The BASB034 products of the invention can be used
CC screen for new antibacterial compounds that may target resistant
CC bacteria.

CC
XX
SQ Sequence 442 AA:

Query Match 100.0%; Score 2360; DB 21; Length 442;

Best Local Similarity 100.0%; Pred. No. 1.3e-218;

Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVSLSTLTLSISCFALIAIOAKAVPNVAFVDEVRESENDLGONDELPIDVOSATQSA 60
DB 1 MKVSLSTLTLSISCFALIAIOAKAVPNVAFVDEVRESENDLGONDELPIDVOSATQSA 60
QY 61 STDANPLDEHEPELYTTALENKTMILNCSALNODIMRLACYDPLVHGEPYAVIKTRRSI 120
DB 61 stdanpldehepe lyttalenktmlncsalnodimrlacydplvngepavlyktrrsi 120
QY 121 RLDEETIMQTIKGRPOVYVOETTPPIFLMGNEKGMILTKKDAKOLEYAKOFTPLSLSDLD 180
DB 121 rldeetimqti kgrpovvyoettp piflmgn ekgmiltkkda koleyakof tplsldld 180
QY 181 RNNTPILWSSRPANPMYULPIFMHGKPNRSPNTPSHEAROPTNFEFRAPELKFOYSVYKA 240
DB 181 rnn tpi lws srp an pmy ul p if mhg k pnr sp ntp s hea ro p t n f e fr a p e l k f o y s v y k a 240
QY 241 AEDLMGTSDLMFGYTTQOSHMOJFNGKNSRPFVNDYQPEIFLTQPYSDLPMDGKVRMI 300
DB 241 aedlmg tsd lmf g y t t q o s h m o j f n g k n s r p f v n d y q p e i f l t q p y s d l p m d g k v r m i 300
QY 301 GNGAVHHNSNGESAKLSRSNRRATILMAGMEKNLTVMPIRGRIFFKESSGQPDNDPIILD 360
DB 301 gngav h h n s n g e s a k l s r s n r r a t i l m a g m e k n l t v m p i r g r i f f k e s s g q p d n d p i i l d 360
QY 361 YYGVDVRFYLYOLEKNSNIGTAVYNRSKGALQLDYVPLGKIGISYQOIFOGYQOSL 420
DB 361 yygvdv r f y l y o l e k n s n i g t a v y n r s k g a l q l d y v p l g k i g i s y q o i f o g y q o s l 420
QY 421 IDYNHEATSFVGILMDMNGL 442
DB 421 idynheatsfvgilm d m n g l 442

RESULT 2
ID AA85268 standard; Protein; 442 AA.
XX
AC AA85268;
XX

DT 29-JUN-2000 (first entry)
XX
XX BASB034 amino acid sequence #1.

KW Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW hearing loss; antibacterial drug.

XX Moraxella catarrhalis.

XX WO200015802-A1.

XX 23-MAR-2000.

XX 14-SEP-1999; 99WO-EP06781.

XX 14-SEP-1998; 98GB-0020002.

XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX Ruelle J;

XX WPI: 2000-271440/23.

XX N-PsDB: AAA10700.

PT Novel BASB034 polynucleotides and polypeptides from Moraxella
PT catarrhalis used to prepare vaccines against bacterial infections -

PS Claim 3; Fig 2; 106pp; English.

CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC strain Mc2931 (ATCC 43617). The invention relates to BASB034 polypeptides
CC from M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2959. The
CC BASB034 polynucleotides and polypeptides may be employed as research
CC reagents and material for the discovery of treatments and diagnostics for
CC diseases, particularly human diseases. They are particularly used to
CC diagnose and treat M. catarrhalis infections. They can be used for
CC diagnosis of disease, staging of disease, or determining response of an
CC infectious organism to drugs. The polynucleotides may be used as a source
CC for hybridization probes, and for screening of genetic mutations,
CC serotype, organism or strain identification, identification of mutations
CC in BASB034 sequences, and as components of arrays which are useful for
CC diagnostic and prognostic purposes. The polypeptides can be used to
CC produce antibodies. The polypeptides can also be used in vaccine
CC formulations, and to identify agonists and antagonists. The polypeptides,
CC antibodies, agonists and antagonists (which are bacteriostatic) are used
CC for the treatment and prevention of diseases such as otitis media in
CC infants and children, pneumonia in elderlies, sinusitis, nosocomial
CC infections and invasive diseases, and chronic otitis media with hearing
CC loss. The polypeptides, agonists and antagonists are also used for
CC screening of antibacterial drugs. The BASB034 products of the invention
CC can be used screen for new antibacterial compounds that may target
CC resistant bacteria.

XX
SQ Sequence 442 AA:

Query Match 99.7%; Score 2352; DB 21; Length 442;

Best Local Similarity 99.3%; Pred. No. 7.6e-218;

Matches 439; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVSLSTLTLSISCFALIAIOAKAVPNVAFVDEVRESENDLGONDELPIDVOSATQSA 60
DB 1 MKVSLSTLTLSISCFALIAIOAKAVPNVAFVDEVRESENDLGONDELPIDVOSATQSA 60
QY 61 STDANPLDEHEPELYTTALENKTMILNCSALNODIMRLACYDPLVHGEPYAVIKTRRSI 120
DB 61 stdanpldehepe lyttalenktmlncsalnodimrlacydplvngepavlyktrrsi 120
QY 121 RLDEETIMQTIKGRPOVYVOETTPPIFLMGNEKGMILTKKDAKOLEYAKOFTPLSLSDLD 180
DB 121 rldeetimqti kgrpovvyoettp piflmgn ekgmiltkkda koleyakof tplsldld 180

```

OY 181 RNNFTPLMSSRRHNPMPYVLPPIFMHGKPNRSPPTPSHEARQPTPSHARAPELKEOYSVKYKA 240
DB 181 rnnprlssrrhnpmpyvlppifmhygkpnrsptpsheakqfipnrtapelkfysvkvka 240
OY 241 AEDLMGTDSDLMFGYTGQSHMOIFNGKNSRPRVHDYQPEIFLTPQVYSDLPMDGKVMIM 300
DB 241 aedltwgtddslwfgtqshwqifngknsrprfvrhdyqpelfltpqpydsdpwgdgkvrm 300
OY 301 GMGAVHSHNGESAKLSRSMNRAIYLMAGMEWKNLTVMPTIRIGRIFKESGSGQDDNDPDL 360
DB 301 gmgavhshngesaklsrsmnraylmagmewknltvmptirwgrlfksgsgsqddndpdlld 360
OY 361 YYGVDVRFVLQLENKNSISGTVRNPRSGKALOLDVYVPLGKISGTFQIFQGYGSL 420
DB 361 yygvgdvrfllqlenknsisgltvrynprsgkqalqldvypdlygkysgyfqlfgygqsl 420
OY 421 IDYNHEATSPFGVGLMLNDMGMGL 442
DB 421 idynheatsfgvglmndmngl 442

RESULT 3
ID AAY85271 standard; Protein: 442 AA.
XX
XX AAY85271:
XX
XX 29-JUN-2000 (first entry)
XX
XX BASB034 amino acid sequence #4.
XX
XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;
XX vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
XX sinusitis; nosocomial infection; invasive disease; chronic otitis media;
XX hearing loss; antibacterial drug.
XX
XX Moraxella catarrhalis.
XX
XX W0200015802-A1.
XX
XX 23-MAR-2000.
XX
XX 14-SEP-1999; 99WO-EP06781.
XX
XX 14-SEP-1999; 98GB-0020002.
XX
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Ruelle J;
XX
XX WPI; 2000-271440/23.
XX
XX N-PSDB; AAA10703.
XX
XX Novel BASB034 polynucleotides and polypeptides from Moraxella
XX catarrhalis used to prepare vaccines against bacterial infections
XX
XX Claim 3; Page 69; 106pp; English.
XX
XX This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
XX strain Mc2969. The invention relates to BASB034 polypeptides from
XX Moraxella catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
XX polynucleotides and polypeptides may be employed as research reagents and
XX material for the discovery of treatments and diagnostics for diseases,
XX particularly human diseases. They are particularly used to diagnose and
XX treat M. catarrhalis infections. They can be used for diagnosis of
XX disease, staging of disease, or determining response of an infectious
XX organism to drugs. The polynucleotides may be used as a source for
XX hybridization probes, and for screening of genetic mutations, serotype,
XX organism or strain identification, identification of mutations in BASB034
XX sequences, and as components of arrays which are useful for diagnostic
XX and prognostic purposes. The polypeptides can be used to produce
XX antibodies. The polypeptides can also be used in vaccine formulations,

```

```

CC and to identify agonists and antagonists. The polypeptides, antibodies,
CC agonists and antagonists (which are bacteriostatic) are used for the
CC treatment and prevention of diseases such as otitis media in infants and
CC children, pneumonia in elderly, sinusitis, nosocomial infections and
CC invasive diseases, and chronic otitis media with hearing loss. The
CC polypeptides, agonists and antagonists are also used for screening of
CC antibacterial drugs. The BASB034 products of the invention can be used
CC screen for new antibacterial compounds that may target resistant
CC bacteria.
XX
XX Sequence 442 AA:
XX
XX
XX Query Match 99.5%; Score 2348; DB 21; Length 442;
XX Best Local Similarity 99.3%; Pred. No. 1.8e-217;
XX Matches 439; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 1 MKVSLSTLTLSILSCFALLAIQAAKAVPNPAFVDEVSENDLGQDNELPIDVQATQSA 60
DB 1 mkvslstltlslpofallaigqagavpnpvafdevrsendlgqdnelpidvqatqsa 60
OY 61 STDANPIDEHEPELTYTALENKTMWLINCSALNDIMFLACYDTLVHGETPAVITKRSI 120
DB 61 stdanpidehepeltytalenktmlncsalngdimrlyacdytlvhgetpaviktksi 120
OY 121 RLDERIMQTIKGRQVYVQETDPIFLMGNEKGMILTKKDAKOLEAAKQFTPLSFDLD 180
DB 121 rldeitwqtkgrqvyvqetdprflmgnegmiltkkdakqleyaakqfplslsfdld 180
OY 181 RNNFTPLMSSRRHNPMPYVLPPIFMHGKPNRSPPTPSHEARQPTPSHARAPELKEOYSVKYKA 240
DB 181 rnnprlssrrhnpmpyvlppifmhygkpnrsptpsheakqfipnrtapelkfysvkvka 240
OY 241 AEDLMGTDSDLMFGYTGQSHMOIFNGKNSRPRVHDYQPEIFLTPQVYSDLPMDGKVMIM 300
DB 241 aedltwgtddslwfgtqshwqifngknsrprfvrhdyqpelfltpqpydsdpwgdgkvrm 300
OY 301 GMGAVHSHNGESAKLSRSMNRAIYLMAGMEWKNLTVMPTIRIGRIFKESGSGQDDNDPDL 360
DB 301 gmgavhshngesaklsrsmnraylmagmewknltvmptirwgrlfksgsgsqddndpdlld 360
OY 361 YYGVDVRFVLQLENKNSISGTVRNPRSGKALOLDVYVPLGKISGTFQIFQGYGSL 420
DB 361 yygvgdvrfllqlenknsisgltvrynprsgkqalqldvypdlygkysgyfqlfgygqsl 420
OY 421 IDYNHEATSPFGVGLMLNDMGMGL 442
DB 421 idynheatsfgvglmndmngl 442

RESULT 4
ID AAY85269 standard; Protein: 442 AA.
XX
XX AAY85269:
XX
XX 29-JUN-2000 (first entry)
XX
XX BASB034 amino acid sequence #2.
XX
XX Moraxella catarrhalis.
XX
XX W0200015802-A1.
XX
XX 23-MAR-2000.
XX
XX 14-SEP-1999; 99WO-EP06781.
XX
XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;
XX vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
XX sinusitis; nosocomial infection; invasive disease; chronic otitis media;
XX hearing loss; antibacterial drug.
XX
XX Moraxella catarrhalis.
XX
XX W0200015802-A1.
XX
XX 23-MAR-2000.
XX
XX 14-SEP-1999; 99WO-EP06781.
XX

```

PR 14-SEP-1998; 98GB-0020002.
 XX (SMK) SMITHKLINE BECHAM BIOLOGICALS.
 PA
 XX
 PI Ruelle J;
 XX
 DR WPI: 2000-271440/23.
 XX N-PSDB: AAA10701.
 XX
 PT Novel BASB034 polynucleotides and polypeptides from Moraxella
 XX catarrhalis used to prepare vaccines against bacterial infections
 PS Claim 3; Page 67; 106pp; English.

CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 CC strain Mc2908. The invention relates to BASB034 polypeptides from
 CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
 CC polynucleotides and polypeptides may be employed as research reagents and
 CC material for the discovery of treatments and diagnostics for diseases,
 CC particularly human diseases. They are particularly used to diagnose and
 CC treat M. catarrhalis infections. They can be used for diagnosis of
 CC disease, staging of disease, or determining response of an infectious
 CC organism to drugs. The polynucleotides may be used as a source for
 CC hybridization probes, and for screening of genetic mutations, serotype,
 CC organism or strain identification, identification of mutations in BASB034
 CC sequences, and as components of arrays which are useful for diagnostic
 CC and prognostic purposes. The polypeptides can be used to produce
 CC antibodies. The polypeptides can also be used in vaccine formulations,
 CC and to identify agonists and antagonists. The polypeptides, antibodies,
 CC agonists and antagonists (which are bacteriostatic) are used for the
 CC treatment and prevention of diseases such as otitis media in infants and
 CC children, pneumonia in elders, sinusitis, nosocomial infections and
 CC invasive diseases, and chronic otitis media with hearing loss. The
 CC polypeptides, agonists and antagonists are also used for screening of
 CC antibacterial drugs. The BASB034 products of the invention can be used
 CC screen for new antibacterial compounds that may target resistant
 CC bacteria.

XX Sequence 442 AA;

Query Match 98.7%; Score 2330; DB 21; Length 442;

Best Local Similarity 98.9%; Pred. No. 1e-215; 3; Indels 0; Gaps 0;

Matches 437; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKVSLTLTLTLSCFAIILAQAKAVNPVAFVDEYRESENDLGDNELPIDVQATQSA 60
 DB 1 MKVSLTLTLTLTLSCFAIILAQAKAVNPVAFVDEYRESENDLGDNELPIDVQATQSA 60
 QY 61 STPTANPLDEHEPELTYTALENKTMALINCSALNODIMRLACYDTLVHGETPAVIKTKRSI 120
 DB 61 STPTANPLDEHEPELTYTALENKTMALINCSALNODIMRLACYDTLVHGETPAVIKTKRSI 120
 QY 121 RLDETTMOTKRGKPOVVOYQOTDPIFMGMEKGMLETKKDKKOLEYAKKOTPTSLSDLD 180
 DB 121 RLDETTMOTKRGKPOVVOYQOTDPIFMGMEKGMLETKKDKKOLEYAKKOTPTSLSDLD 180
 QY 181 RNNTPPLMSSRPNNPMVLPFLFMHGNKPNRSPNTPESHEARQPTNFEFRAPDEKFCQSVYKA 240
 DB 181 RNNTPPLMSSRPNNPMVLPFLFMHGNKPNRSPNTPESHEARQPTNFEFRAPDEKFCQSVYKA 240
 QY 241 AEDLWGTDSLWFGYTOQSHWQJFNKNSRPFRRVNDYQPEIFLTQPYSDLPMDGKVRMT 300
 DB 241 AEDLWGTDSLWFGYTOQSHWQJFNKNSRPFRRVNDYQPEIFLTQPYSDLPMDGKVRMT 300
 QY 301 GMGAVHHNSGESAKLSNRRAYLMAGMEKKNLTVMRIRGRIFKEGSGOPDPNDPILD 360
 DB 301 GMGAVHHNSGESAKLSNRRAYLMAGMEKKNLTVMRIRGRIFKEGSGOPDPNDPILD 360
 QY 361 YVGYGVDFLYQLQENKSNISGTVARNRSGKALQLDLYVPLGKGISGYQIRFGYQOSTL 420
 DB 361 YVGYGVDFLYQLQENKSNISGTVARNRSGKALQLDLYVPLGKGISGYQIRFGYQOSTL 420

QY 421 IDYNHEATSGVGLIMLMDMGL 442
 DB 421 IDYNHEATSGVGLIMLMDMGL 442

RESULT 5

AAV75156 standard; Protein; 370 AA.

AAV75156;

21-MAR-2000 (first entry)

Neisseria gonorrhoeae ORF 582 protein sequence SEQ ID NO:1786.

Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;

antibacterial; diagnosis; immunogenic; infection; meningitis; septicemia;

Neisseria gonorrhoeae.

W09957280-A2.

11-NOV-1999.

30-APR-1999; 99MO-US09346.

01-MAY-1998; 98US-0083758.

31-JUL-1998; 98US-0094869.

02-SEP-1998; 98US-0098894.

02-SEP-1998; 98US-0099062.

09-OCT-1998; 98US-0103749.

09-OCT-1998; 98US-0103794.

09-OCT-1998; 98US-0103796.

25-FEB-1999; 99US-0121528.

(CHIR) CHIRON CORP.

(GENO-) INST GENOMIC RES.

Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M;

Petersen J, Piazza M, Rappoli R, Ratti G, Scalato E, Scarselli M;

Tettelin H, Venter JC.

WPI: 2000-062150/05.

N-PSDB; AA253918.

Novel Neisserial polypeptides predicted to be useful antigens for

vaccines and diagnostics

Claim 2; Page 903; 1453pp; English.

AA253015 to AA254536, AA254577 to AA254615, and AAV74253 to AAV75941

represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides

PCR primers used in the exemplification of the present invention. The

polypeptides, the polynucleotides, antibodies and compositions of

the invention can be used as vaccines, as diagnostic reagents, and as

immunogenic compositions. The polypeptides can be used in the

manufacture of medicaments for treating or preventing infection due to

Neisserial bacteria (e.g. meningitis and septicemia), to detect the

presence of Neisseria bacteria, or to raise antibodies. They may also

be used to screen for agonists or antagonists, which may themselves

have use as antibacterial agents. The polynucleotides of the invention

may also be used in gene therapy protocols.

Sequence 370 AA;

Query Match 34.8%; Score 822; DB 21; Length 370;

Best Local Similarity 44.7%; Pred. No. 1.4e-70; Matches 163; Conservative 68; Mismatches 110; Indels 24; Gaps 8;

87 INCSALNODIMRLACYDTLVHGETPAVI-----KTKRSINLDETTIOTI-KGRPOVVOYE 140


```

Db      21  lqcaalldnvtrrlacvdrifaaqlpsasagqegskavlnltetrssldkgeavivvek 80
Qy      141 TTDPIFLMGNEKGMFLTKKADQLEYAAKQFTPLSLSPDLDRNN-TPLMSSRPHPMYLP 199
Db      81  ggdaL-----padsagetadlycplslmyldkndrlgllgvrehnmpmylmp 127
Qy      200 IFMGKPNRSPNTPSHEAR-QFTPNFRAPRLKFOYSKVVAADDMCTSDMLNFGYQO 258
Db      128 fwmynspnyapsprltvyeqkfgqkraetklyvsfkksiaenllfkttradlwfgytqr 187
Qy      259 SHMOIFN-GKNSRPRFVHDYQPEIFLTQPVYSDLPMDCKVMIGMGAHHSNGESAKLSR 317
Db      188 sdwqlyngqrsapfrntdykpeflfcqpvkadlprfgtrlmigaqfghsgngqsrpre 247
Qy      318 SWNRAYLMAGMEKMLTVMPRIWGRIFKEGSGSQPDNDPDLIDYGYCDVDFLYOLENKS 377
Db      248 swntlymagmewgklvlpvrvvrafdq-sgdk-ndnppdladyngydvklygrylndrq 305
Qy      378 NISGTVRYNPRSGKALQLDVYVPLGKISGFQIFOGYGSGLDYNHEATSPFGVGLMLN 437
Db      306 nvysvlyrnpktygaleaaytfplkqklygvrvfngygeslidyhkhqngiglylfn 365
Qy      438 DMWGL 442
Db      366 dwdgi 370

RESULT 6
AA75157
ID  AA75157 standard; Protein: 370 AA.
AC  AA75157;
XX  21-MAR-2000 (first entry)
DE  Neisseria meningitidis ORF 582 protein sequence SEQ ID NO:1788.
XX  XX
KW  Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW  antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW  antibacterial; gene therapy.
XX  XX
OS  Neisseria meningitidis.
XX  XX
PN  W09957280-A2.
XX  PD
XX  11-NOV-1999.
XX  PE
XX  30-APR-1999; 99WO-US09346.
XX  PR
XX  01-MAY-1998; 98US-0083758.
XX  PR  31-JUL-1998; 98US-0094869.
XX  PR  02-SEP-1998; 98US-0098994.
XX  PR  02-SEP-1998; 98US-0099062.
XX  PR  09-OCT-1998; 98US-0103749.
XX  PR  09-OCT-1998; 98US-0103794.
XX  PR  09-OCT-1998; 98US-0103796.
XX  PR  25-FEB-1999; 99US-0121528.
XX  XX
PA  (CHIR ) CHIRON CORP.
PA  (GENO-) INST GENOMIC RES.
XX  XX
PI  Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M,
PI  Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M,
PI  Tettelin H, Venter JC.
XX  XX
DR  WPI: 2000-062150/05.
DR  N-PSDB: AA253919.
XX  XX
PT  Novel Neisserial polypeptides predicted to be useful antigens for
PT  vaccines and diagnostics
XX  XX
PS  Claim 2; Page 904; 1453pp; English.

```

```

XX  CC  AA253015 to AA254536, AA254577 to AA254615, and AA74253 to AA75941
XX  CC  represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
XX  CC  and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
XX  CC  PCR primers used in the exemplification of the present invention. The
XX  CC  polynucleotides, the polynucleotides, antibodies and compositions of
XX  CC  the invention can be used as vaccines, as diagnostic reagents, and as
XX  CC  immunogenic compositions. The polypeptides can be used in the
XX  CC  manufacture of medicaments for treating or preventing infection due to
XX  CC  Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
XX  CC  presence of Neisseria bacteria, or to raise antibodies. They may also
XX  CC  have use as antigen for agonists or antagonists, which may themselves
XX  CC  may also be used in gene therapy protocols.
XX  XX
XX  Sequence 370 AA;
XX  XX
Query Match 34.5%; Score 815; DB 21; Length 370;
Best Local Similarity 44.7%; Pred. No. 6.7e-70;
Matches 163; Conservative 67; Mismatches 111; Indels 24; Gaps 8;
Qy      87  INCSALNODIMRLACVDTLVHGETPAVI-----KTKRSIRLDETIWQT-RGKPOVYQOE 140
Db      21  lqcaalldnvtrrlacvdrifaaqlpsasagqegskavlnltetrssldkgeavivvek 80
Qy      141 TTDPIFLMGNEKGMFLTKKADQLEYAAKQFTPLSLSPDLDRNN-TPLMSSRPHPMYLP 199
Db      81  ggdaL-----padsagetadlycplslmyldkndrlgllgvrehnmpmylmp 127
Qy      200 IFMGKPNRSPNTPSHEAR-QFTPNFRAPRLKFOYSKVVAADDMCTSDMLNFGYQO 258
Db      128 fwmynspnyapsprltvyeqkfgqkraetklyvsfkksiaenllfkttradlwfgytqr 187
Qy      259 SHMOIFN-GKNSRPRFVHDYQPEIFLTQPVYSDLPMDCKVMIGMGAHHSNGESAKLSR 317
Db      188 sdwqlyngqrsapfrntdykpeflfcqpvkadlprfgtrlmigaqfghsgngqsrpre 247
Qy      318 SWNRAYLMAGMEKMLTVMPRIWGRIFKEGSGSQPDNDPDLIDYGYCDVDFLYOLENKS 377
Db      248 swntlymagmewgklvlpvrvvrafdq-sgdk-ndnppdladyngydvklygrylndrq 305
Qy      378 NISGTVRYNPRSGKALQLDVYVPLGKISGFQIFOGYGSGLDYNHEATSPFGVGLMLN 437
Db      306 nvysvlyrnpktygaleaaytfplkqklygvrvfngygeslidyhkhqngiglylfn 365
Qy      438 DMWGL 442
Db      366 dldgi 370

RESULT 7
AA75158
ID  AA75158 standard; Protein: 370 AA.
AC  AA75158;
XX  XX
XX  21-MAR-2000 (first entry)
DE  Neisseria meningitidis ORF 582 protein sequence SEQ ID NO:1790.
XX  XX
KW  Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW  antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW  antibacterial; gene therapy.
XX  XX
OS  Neisseria meningitidis.
XX  XX
PN  W09957280-A2.
XX  PD
XX  11-NOV-1999.
XX  PE
XX  30-APR-1999; 99WO-US09346.
XX  XX

```


QY 217 ARQFTPNEFRAPELKFOYVSVKKAEDLMGTDSDLMFGYTOOSHQIENGKNSRPFVRVD 276
Db 103 ---ygrnef-----kfqlsfvrvfthilwtkcllylaqtcdwfgdyndpqsapmrimn 154
QY 277 YQPEIFLTPQPVYSDLPWDKGV---RMIGGAVHNSNG--ESAKLSNSMNRAVYLMAGMEKN 332
Db 155 fmpelllyvypI-nfkpfqgkIqnfseIwqghlsngvgagcyqpfnk-----egnpeng 209
QY 333 LTVMPRI-----WGRIFKEGSGSQP-----DDNPD 357
Db 210 fpgqvnIvkdyngqkdvrgwgcsvsagqrpvfrIwwekggIkmwavywpyvpydqsnpn 269
QY 338 ILDYGYGVDVRFly-----OLENKSNIsgTVRYRNPRSGGALQLDVYVPLGKISGYF 410
Db 270 IldymgygnakIdyrrgrhfeIdqIdftgywryd--rwbgafrIgytyrInpfvgyIa 327
QY 411 QIFOGYGOSLIDYNHEATSFVGLMLN 437
Db 328 qwfnsgygdqIyeydvfnrIgyvgrIn 354

RESULT 11

AA10960

ID AA10960 standard; Protein: 356 AA.

XX AC AAY10960;

DT 08-JUN-1999 (first entry)

DE H. pylori ORF 07ap80601_5083193_f3_8 cell envelope protein.

KW Vaccine; probe: diagnostic; ORF; cell envelope protein;

KM secreted protein; cellular protein.

OS Helicobacter pylori.

XX PN W09818323-A1.

PD 07-MAY-1998.

PF 28-OCT-1997; 97WO-US19575.

XX PR 14-JUL-1997; 97US-0891928.

PR 28-OCT-1996; 96US-0739150.

PR 06-DEC-1996; 96US-0759739.

XX PA (ASTR) ASTRA AB.

PI Alm RA, Smith D;

XX DR WPI; 1998-271811/24.

DR N-PSDB; AAX30427.

XX PT Helicobacter pylori nucleic acids and proteins - used to develop

PT products for the detection, prevention and treatment of H. pylori

PT infections

XX PS Claims 27, 31; Page 160-161; 279pp; English.

XX CC Recombinant or substantially pure preparations of H. pylori polypeptides

CC are disclosed, together with the nucleic acids encoding them. In all,

CC 73 ORFs are shown. The proteins are variously cell envelope proteins,

CC secreted proteins or other cellular proteins. Vaccines containing the

CC nucleic acids or proteins are claimed, as are probes containing at least

CC 8 nucleotides from the nucleic acid sequences. The vaccines are useful

CC for treating or reducing the risk of H. pylori infections, and the

CC probes can be used diagnostically for detecting the presence of

CC Helicobacter in a sample. The products are also of use in screening

CC for compounds having the ability to interfere with the H. pylori life

CC cycle or to inhibit H. pylori infection.

XX Sequence 356 AA;

Query Match 10.1%; Score 239.5; DB 19; Length 356;
Best Local Similarity 24.8%; Pred. No. 1.7e-14;
Matches 82; Conservative 43; Mismatches 111; Indels 95; Gaps 12;

QY 157 KDAQOLEYAAKQFTPLSEFDLRNNRPLWSSRRHNPVYVPIRMHGKPNRSPPTPSHE 216
Db 70 kKylmmmdyIgyfIPIYhsf-----cpIIdqwyhpnInp----- 103
QY 217 ARQFTPNEFRAPELKFOYVSVKKAEDLMGTDSDLMFGYTOOSHQIENGKNSRPFVRVD 276
Db 104 ---ygrnef-----kfqlsfvrvfthilwtkcllylaqtcdwfgdyndpqsapmrimn 155
QY 277 YQPEIFLTPQPVYSDLPWDKGV---RMIGGAVHNSNG--ESAKLSNSMNRAVYLMAGMEKN 332
Db 156 fmpelllyvypI-nfkpfqgkIqnfseIwqghlsngvgagcyqpfnk-----egnpeng 214
QY 323 -----YLMAGMEKN--NLTVPRIWGRIRKESGSGOPDD- 354
Db 215 pvlvkdynqgkdvrgwgcsvsagqrpvfrIwwekggIkmwavywpyvpydqsnpn 269
QY 355 -NPDILDYGYGVDVRFly-----OLENKSNIsgTVRYRNPRSGGALQLDVYVPLGKISGYF 406
Db 267 snpqIldymgygnakIdyrrgrhfeIdqIdftgywryd--rwbgafrIgytyrInpfv 324
QY 407 SGYFOIFOGYGOSLIDYNHEATSFVGLMLN 437
Db 325 gIyagqwfngygdqIyeydvfnrIgyvgrIn 355

RESULT 12

AAW20760

ID AAW20760 standard; Protein: 253 AA.

XX AC AAW20760;

DT 15-JUL-1997 (first entry)

DE H. pylori outer membrane protein, 07ap80601orf8.

KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;

KM binding compound; bacterium; life cycle; activator; bacteria; inhibitor;

KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope;

XX OS Helicobacter pylori.

XX PN W09640893-A1.

PD 19-DEC-1996.

PF 06-JUN-1996; 96WO-US09122.

XX PR 01-APR-1996; 96US-0630405.

PR 07-JUN-1995; 95US-0487032.

XX PA (ASTR) ASTRA AB.

PI Berglindh OT, Smith D, Mellgaerd BL;

XX DR WPI; 1997-052306/05.

DR N-PSDB; AAT68013.

XX PT Helicobacter pylori nucleic acid sequences and related

PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori

PT infection, and to detect Helicobacter

XX PS Claim 56; Page 1172-1173; 1481pp; English.

XX CC The present sequence is a Helicobacter pylori outer membrane protein.

CC The protein may be used in a vaccine to prevent or treat H. pylori

CC infection or to identify H. pylori polypeptide binding compounds, useful

CC as potential H. pylori life cycle activators or inhibitors. The genomic

sequence of H. pylori (ATCC 55679) was determined from overlapping cDNAs generated by mechanically shearing the bacterial DNA. The sequences were analysed for ORF of at least 180 nucleotides, and the predicted coding regions defined by computer evaluation. To identify likely H. pylori antigens for vaccine development, the amino acid sequences predicted from various ORF were analysed for significant homology to other known or exported membrane proteins. Having identified and determined the sequences of interest, particular regions can be isolated from H. pylori by PCR amplification for recombinant polypeptide production, e.g. in *E. coli* hosts.

SQ Sequence 253 AA;

Query Match	6.68;	Score 156.5;	DB 18;	Length 253;
Best Local Similarity	25.68;	Pred. No. 1e-06;		
Matches 56;	Conservative 26;	Mismatches 72;	Indels 65;	Gaps 8

[illegible]

RESULT 13
2A470001

ID AAW70991 standard; 1686 AA.

AC AAW70991;

DT 19-OCT-1998 (first entry)

DE Human class II P13 kinase-C2alpha.

KW Human; class II phosphoinositide lipid kinase; PI3 kinase

XX

XX 3

XX
XX
77-40070000

30-JUL-1998

2/-JAN-1998; 98WO-GB00244

PR 28-JAN-1997; 97GB-0001652.

PA (LUDW-) LUDWIG INST CANCER RES.

PI Domin J, Waterfield MD;

DR WPI: 1998-427960/36.

[illegible]

PT fragments - useful for e.g. treatment of tumour cells where

PI phenotype is associated with expression of kinase

Claim 3; Fig 1; 52pp; English.

CC The present sequence represents

CC kinase due to the presence of a conserved C2 domain found in murine and
CC Drosophila class II p13 kinases, its apparent lack of a p85 binding site
CC and a substrate affinity to inositol lipids PtdIns and PtdIns(4)P. The
CC protein has resistance to p13 kinase inhibitors Wortmannin and LY294002.
CC Antibodies against the protein (optionally humanised), are used to
CC identify class II p13 kinases. Antisense sequences, antibodies or
CC dominant negative mutants of the p13-C2alpha protein, are useful in human
CC or veterinary medicine to block class II kinases. They can be used to
CC treat tumour cells where the phenotype is associated with expression of
CC p13-C2alpha protein.

Sequence 1686 AA; SQ

Query Match	4.8%	Score 114	DB 19	Length 1686
Best Local Similarity	19.0%	Pred. No. 0.25		
Matches 85	Conservative 68	Mismatches 154	Indels 140	Gaps 21

```

OY 45 QDNELPIDV-----QSATQASSTPTAPRLDE-----HEPELYTTALENKMMLN----- 88
Db 522 eddeprvdlnhlygiekpckeamtrpveellidsynqvelalqtenjhradvqkav 581
OY 89 ---CSALNODIMRLACYDYLHCGETPNAVITKRSIRLDEIWMOTIKGKPOVVUOETTPDI 145
Db 582 rkicsald-gyettai-----tesvkkikravmlpr-----kredvc 618
OY 146 FLMGNEKMLTKKAKOLEVAAKOFPLUSFDLDNRNNTPLMSSRRPNEMTV-----LP 199
Db 619 slfge-----dixrest-rgslnpenpvyasinqtae 651
OY 200 IF-----MHGKRNRSPTNPSHEAROFTEPNEFRAPLEKQVSVKVAABEDL---WGTDSDLW 252
Db 652 lydlirrhansgreptcdagsgskveawtltqgltf-----aahgisnwnvsnky 707
OY 253 FGYTQOSHWOLFNGKN-SRPF---RVHDYOELFLTO-----PV-YSDLPMDGKVRMT 300
Db 708 ylicslsh---ngkdlfkipqskkvtyknfyllfwdeliiflpiqslplessvhl 753
OY 301 GMGAVVHSHNGESALTSRSMN-----RAYLWAGMEWMKULGYMPRIWGRIFK 345
Db 764 lfgalingssgspsobnkqyrkyprealgkvsjpldcfrflcgy-----tlllylw 812
OY 346 EGSGSQPDNDPDLIDYVGGDVREFLYOLENKSNSITGVRRYRNPBSGKALOLDVYVPJGK 405
Db 813 -teshtnsygvttkkgymeriavlqvdffspafaillytptydvrsilqhmletend 870
OY 406 ISGFYQIFQSGOSLIDYNEHNTSFV 432
Db 871 lkg-----kldllhkdsisl 887

```

RESULT 14

ID AAP70420 standard; protein; 800 AA.

AC AAP70420

DT 20-JAN-1991 (first entry)

DE Sequence encoded by cellulase gene derived from *Bacillus* sp. No. 1139.

KW Enzyme; celotriose; celotetrose; hydrolysis

Bacillus sp. No. 1139.

AND

ET	Peptide	1:30
31	800	

XX
XX
TP6333296-A

XX
X

XX

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:17:45 ; Search time 36.79 Seconds
(Without alignments)
270.358 Million cell updates/sec

Title: US-09-787-083-6

Perfect score: 2360
Sequence: 1 MKVSLSTLTSLSCFAILA.....YNHEATSGVGLMDMGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
- 2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
- 3: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
- 4: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
- 5: /cgn2_6/prodata/2/1aa/PCBUS_COMB.pep:*
- 6: /cgn2_6/prodata/2/1aa/backfilseq1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100.5	4.3	682	3	US-08-481-435-6
2	100.5	4.3	1726	2	US-08-609-049A-30
3	100.5	4.3	1726	4	US-09-170-996-30
4	98.5	4.2	1658	2	US-08-609-049A-13
5	98.5	4.2	1658	4	US-09-170-996-13
6	90.5	3.8	1024	4	US-09-091-117-5
7	89.5	3.8	857	2	US-07-717-331E-2
8	88	3.7	537	2	US-08-633-879C-2
9	87.5	3.7	503	1	US-07-946-497-2
10	87.5	3.7	503	2	US-08-483-322-2
11	87.5	3.7	503	2	US-08-478-882-2
12	87.5	3.7	666	4	US-08-961-083-2
13	86	3.6	781	4	US-08-373-134D-2
14	86	3.6	781	2	US-09-114-637-2
15	86	3.6	1088	3	US-08-633-768A-1
16	85	3.6	816	1	US-07-731-157A-4
17	85	3.6	816	1	US-08-229-444B-2
18	85	3.6	816	2	US-08-541-780-4
19	85	3.6	985	5	PCR-US96-03916-6
20	85	3.6	985	5	PCR-US96-03916-66
21	84.5	3.6	774	3	US-08-802-632-2
22	84.5	3.6	774	3	US-09-073-354-1
23	84.5	3.6	774	3	US-08-656-005A-1
24	84.5	3.6	774	4	US-09-073-259-1
25	84.5	3.6	774	4	US-09-363-095-1
26	84.5	3.6	774	4	US-09-418-027-1
27	84	3.6	522	6	RE34606-6

28	84	3.6	1627	1	US-07-665-792E-9	Sequence 9, App1
29	83	3.5	320	2	US-08-245-511-4	Sequence 4, App1
30	83	3.5	320	2	US-08-600-993A-4	Sequence 4, App1
31	82.5	3.5	657	4	US-09-306-593-2	Sequence 2, App1
32	82.5	3.5	797	4	US-09-086-912-2	Sequence 2, App1
33	82	3.5	527	2	US-08-592-126-145	Sequence 145, App
34	82	3.5	527	2	US-08-687-080-48	Sequence 48, App
35	82	3.5	535	2	US-08-633-879C-4	Sequence 4, App1
36	81.5	3.5	663	1	US-08-441-139-7	Sequence 7, App1
37	81.5	3.5	844	2	US-07-731-157A-6	Sequence 6, App1
38	81.5	3.5	844	2	US-08-541-780-6	Sequence 6, App1
39	81	3.4	355	1	US-07-946-497-5	Sequence 5, App1
40	81	3.4	355	2	US-08-483-322-5	Sequence 5, App1
41	81	3.4	355	2	US-08-478-882-5	Sequence 5, App1
42	81	3.4	1121	1	US-07-789-915A-2	Sequence 2, App1
43	81	3.4	1121	1	US-08-005-002C-2	Sequence 2, App1
44	81	3.4	1121	1	US-08-487-203A-2	Sequence 2, App1
45	80	3.4	334	2	US-08-359-850-4	Sequence 4, App1

ALIGNMENTS

RESULT 1
US-08-481-435-6
Sequence 6, Application US/08481435
Patent No. 6027906
GENERAL INFORMATION:
APPLICANT: Balganes, Tanjore S
TITLE OF INVENTION: No. 6027906el Polypeptides
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESSES:
ADDRESSEE: White & Case
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,435
FILING DATE: 10-JUL-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IN 580/MAS/94
FILING DATE: 01-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9404072-2
FILING DATE: 24-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Steiner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-151
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 682 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-481-435-6

Query Match 4.3%; Score 100.5; DB 3; Length 682;
Best Local Similarity 20.4%; Pred. No. 0.11;
Matches 92; Conservative 68; Mismatches 189; Indels 103; Gaps 24;

```

QY 13 LSCFAILLIOQAKVAVPVVAFVDEVRSENDL-GQDNELPIDVQSAOS-----AST 62
Db 5 LSESEIVATTSKITYDKNKNOLLADLGBERRRMAQANDIPIDVKAIVYSIEDHREFDHRI 64
QY 63 DTANPDLDEHEBELYTTLALENKTMLNCSALNODIMRLACYOTLVHGSETPAVYIKTKRSIRL 122
Db 65 DTRILGAFELRNLOSNSISOG-----SALTOOQIKILTYESTSDOTIS-----RK 110
QY 123 DETIMOTKGRPOVYVOET-----TDPIFJMGNEKSMGLTK-----KRAKOLE----- 164
Db 111 AQEAMLAIOLEOKTKQCEILTYIYNKYMSNGANTGMOTAAONTYKGLNNLSLPOLALLA 170
QY 165 --YAAGOTPLS-LSEFDLDRNNTPLMSSRPHNEMUYLPIFMGKPNRSPTBSHEARQ 220
Db 171 GMPQAPNQUDPYSHPEAAQDRRLVL--SEMKNGYI----SABQYEKAVNTPITDELQS 224
QY 221 TPNEFRPE-----LKTQVYS-VYKKAADLMDGTSDDLMFGYTQ--QSH-WQIFNGKSRP 271
Db 225 LKSASNPAYMDNYLKEIYNQVEEETGYNLTTTMDIYTVNDQEAQKHNMDIYTDVEYA 284
QY 272 FRVHDYQREILTOPYVSDLEPMDCKVRNMGIAVHNHSN-----GESAKLSRSNNRAYLM 325
Db 285 YRPDELO-----VASTIYDVS-NKRY-IAOLGANNQSSNYSFGINQAVETNRDM----- 333
QY 326 AGMEKMKNTL-VMPIRWGRIKEGSGSGSDDDNDPDLIDYGG-----YGDVRELYQ 373
Db 332 -GSTMKRPTDYAPALEVGVY-ESTATTIVHDER--YNYRPGTWTVPYUWMDRGFGYNTLOVA 387
QY 373 LENSNSISGTVRYN-----PRSGKALOYLD 398
Db 388 LQOSRNVPAVETLKNVGLNRAKTYLNGSGIDY 419

```

```

1      RESULT      2
2      US-08-609-049A-30
3      : Sequence 30 Application US/08609049A
4      Patent No. 5948664
5      GENERAL INFORMATION:
6      APPLICANT: Williams, Lewis T.
7      APPLICANT: Molz, Lisa
8      APPLICANT: Chen, Yen-Men
9      TITLE OF INVENTION: NO 5948664el PI 3-Kinase Polypeptides
10     NUMBER OF SEQUENCES: 32
11     CORRESPONDENCE ADDRESS:
12     ADDRESSEE: Townsend and Townsend and Crew LLP
13     STREET: Two Embarcadero Center, 8th Floor
14     CITY: San Francisco
15     STATE: California
16     COUNTRY: USA
17     ZIP: 94111-3834
18     COMPUTER READABLE FORM:
19     MEDIUM TYPE: Floppy disk
20     COMPUTER: IBM PC compatible
21     OPERATING SYSTEM: PC-DOS/MS-DOS
22     SOFTWARE: PatentIn Release #1.0, Version #1.25
23     CURRENT APPLICATION DATA:
24     APPLICATION NUMBER: US/08/609,049A
25     FILING DATE: 29-FEB-1996
26     CLASSIFICATION: 435
27     ATTORNEY/AGENT INFORMATION:
28     NAME: Dow, Karen B.
29     REGISTRATION NUMBER: 29,684
30     REFERENCE/DOCKET NUMBER: 2307K-063700US
31     TELECOMMUNICATION INFORMATION:
32     TELEPHONE: 415-326-2400
33     TELEFAX: 415-326-2422
34     INFORMATION FOR SEQ ID NO: 30:
35     SEQUENCE CHARACTERISTICS:
36     LENGTH: 1726 amino acids
37     TYPE: amino acid
38     TOPOLOGY: linear
39     MOLECULE TYPE: protein

```

US-08-609-049A-30

Query Match	4.38;	Score 100.5;	DB 2;	Length 1726;
Best Local Similarity	18.68;	Pred. No. 0.54;		
Matches 98;	Conservative 79;	Mismatches 190;	Indels 161;	Gaps 21;

QY	1	MKXSL-----TLTSLISCALIAIOAAVPRVAVFV-----EVRENDGOD	46
Db	465	VKVASIEIEGQLPRTFFICDVSFVEIITIMQALSVHDLQNDVGVSLKVCGEVELQ	524
QY	47	NELPIDVQASQASSTJANPDLNEHEPELXTTALENKTALINCSALNODIMRLACTDYL	106
Db	525	NHCGSHENIONCKRWPT-----EIKLOLTLSLMACONLARTAEDD---	565
QY	107	HGETPAVITKRSIRLEDETIMOTIKGRQVVOGETDPIF-----	146
Db	566	--EAP-----VDLKKYLQIQIKRPEKEMTRIRPVELLDSTHYQVELALQTEQNHRAV	615
QY	147	-----LMGNEKGMILTKKAKOLEYA-----KQFRLSISFDLBRNNTPMS	188
Db	616	DOVIAKAVKICISALDGETSPVTEA--VKKLRVNLPRNKSADVTLSGSDTKRKNKGS	674
QY	189	SRHPNPIVUP-----IFMHGKPNR-----SPMTPSEAROFNPEFARDELKFOV	234
Db	675	LNPENPQVMSDHLTTITTYDLRLHANSKCSICGCPGSGNICEAWATE-----QL	726
QY	235	SVKKAEDL--WGTDSDLMFGYTOOSHMOIFNGKN--SRP--RVNDYQREIFLQ--	285
Db	727	QFTYAAHAGISSMNVSYEKUYLLCSLH---NGKDLFPKIOSKVGTYKKNFEYLIKMD	782
QY	286	-----PV-YSDLPMDGKVRMIGMAVHNSGEASAKLSRNN-----RAYL	324
Db	783	ELIIFPIQIOSLPRESVHLHLEFVLNQSSGSSPDSKOKKGBEALGKSVLTFDKRFL	842
QY	325	MAGMEKNTLVMPRIWIRIFKEGSGSQDDNPDLIDYGYGADVRFYOLENKNSIGTVR	384
Db	843	TCG-----FKLELW-----TSSHNSINPCAIIPKKSVMERIVLOVDFSPAPDIY	889
QY	385	YNPRSGALQLODVYVPLGSGISCFPIFGQYOGSLIDYVHNHMAEPRCV	432
Db	890	TSPIDIRNIIOODKLETFLESDIK-----KLDITIHRSDFCL	927

RESULT 3
 US-09-170-996-30
 : Sequence 30, Application US/09170996
 : Patent No. 6291220
 : GENERAL INFORMATION:
 : APPLICANT: Williams, Lewis T.
 : APPLICANT: Molz, Lisa
 : APPLICANT: Chen, Yen-Wen
 : TITLE OF INVENTION: No. 6291220el PI 3-Kinase Polypeptides
 : NUMBER OF SEQUENCES: 32
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Townsend and Townsend and Crew LLP
 : STREET: Two Embarcadero Center, 8th Floor
 : CITY: San Francisco
 : STATE: California
 : COUNTRY: USA
 : ZIP: 94111-3834
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/170,996
 : FILING DATE:
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/609,049
 : FILING DATE: 29-FEB-1996


```

;
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-0637000S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ. ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1726 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-170-996-30

Query Match 4.3%; Score 100.5; DB 4; Length 1726;
Best Local Similarity 18.6%; Pred. No. 0.54;
Matches 98; Conservative 79; Mismatches 190; Indels 161; Gaps 21;

QY 1 MKVSL-----TLTSLSCFAIIAIOAKAVNPVAFV-----EYRSENDIGD 46
DB 465 VKVSIIEGLQLPVTECDVSTVEIIMQALSWHDLQVDVGYTLKVGQEEVLQN 524
QY 47 NELPIDVQASSTDPANPLDEHEPELTTALENKTMLINCSALNODIMRLACVDTLV 106
DB 525 NHCIGSHENIQNCKKMDT-----ETKQLTLTUSAMCONIARTRADD--- 565
QY 107 HGETPAVTKTRSIIRDETIMOTIKGRQVVOETTPIF----- 146
DB 566 --EAP-----VDLKNLYQIEKPKYKEMTRHPVELLDSDYHYQVELALQFENQRAV 615
QY 147 -----LMGNEKMLTKDKAKOLEVAA-----KQTPPLSLSDLDLNNPLWS 188
DB 616 DOYIKAVRKICSLDGVETPSTVEA--VKKLKRAVNLPRNKSADVTSLSGSDTRKNSYKGS 674
QY 189 SRPHNPVYLP-----IFMKGKPR-----SPNPSHEARQTPNPEPAPELQY 234
DB 675 LNPENPVQVSMHLTRITRYDLRLHANSRCSGTCGPRGSRNKEAWTATE-----QL 726
QY 235 SVKVAEDL---WGTDSDLMPGYTOOSHMOIFNGKN--SRPF---RVHDIQPEIFLQ-- 285
DB 727 QLFVYAAHGSISSNMVSNYEKYLLCSLSH---NGKDLFRPIQSKKATGYNKFFYLLKMD 782
QY 286 ----PV-YSDLPWDGKVMIGMGAHVHNSGESAKLSRSWN-----RAYL 324
DB 783 ELIIFPIQSLPLESVLHLLTFLGVLLNOSGSSPDSNKKORKEPALGKVSILTFDFKRL 842
QY 325 MAGMEKNLTVMRIRIRIKESGSGOPDNDPILDYGGVDFRLYQLENKSNISGTVR 384
DB 843 TCG-----TKLLYLW-----TSSHTNSIPGAIIPKSYVMERIVLQVDFPSPADITY 889
QY 385 YNPRSGKALQLDVYVPLGKIGSGYFOIFOGYGSOLIDVNHGATSGV 432
DB 890 TSPQIDRNIIQDKLETLESIDKG-----KLLDIHRDSSGFL 927

RESULT 4
US-08-609-049A-13
; Sequence 13, Application US/08609049A
; Patent No. 5948664
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Moliz, Lisa
; APPLICANT: Chen, Yen-wen
; TITLE OF INVENTION: No. 5948664el PI 3-Kinase Polypeptides
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
```

```

;
; ZIP: 94111-3834
; COMPUTER READABLE FORM.
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,049A
; FILING DATE: 29-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-0637000S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ. ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1658 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-609-049A-13

Query Match 4.2%; Score 98.5; DB 2; Length 1658;
Best Local Similarity 18.9%; Pred. No. 0.81;
Matches 101; Conservative 78; Mismatches 195; Indels 159; Gaps 23;

QY 5 LSTLTSLSCFAIIAIO--QAKAVNPVAFVDEVS-----ENDLGONE 48
DB 381 LSPVTYVRNMGENASVKSIEIGQLPVTFCDVSSVEIIMQALCWVHDLNQ-- 437
QY 49 LPIDVOS-----ATQASSTDPANPLDEHE--PELTTALENKTMLINCSALNODIMRLAC 101
DB 438 --VDVGSITLKVGQEEVLQNNHCLGSHENIQNCKKMDTEIKQLTLTUSAMCONIARTRAE 495
QY 102 YDPLVHGETPAVTKTRSIIRDETIMOTIKGRQVVOETTPIF----- 146
DB 496 DD-----EAP-----VDLKNLYQIEKPKYKEMTRHPVELLDSDYHYQVELALQFENQRAV 542
QY 147 -----LMGNEKMLTKDKAKOLEVAA-----KQTPPLSLSDLDLNNPLWS 183
DB 543 QHRAVDVYKAVRKICSLDGVETPSTVEA--VKKLKRAVNLPRNKSADVTSLSGSDTRKNSYKGS 601
QY 184 TPLMSSRPHNPVYLP-----IFMKGKPR-----SPNPSHEARQTPNPEPAPELQY 229
DB 602 STKGLNPENPVQVSMHLTTATITDRLHANSRCSGTCGPRGSRNKEAWTATE----- 656
QY 230 LKEQVSVKVAEDL---WGTDSDLMPGYTOOSHMOIFNGKN--SRPF---RVHDIQPEIF 282
DB 657 ---QLQFTVYAAHGSISSNMVSNYEKYLLCSLSH---NGKDLFRPIQSKKATGYNKFFY 709
QY 283 LQO-----PV-YSDLPWDGKVMIGMGAHVHNSGESAKLSRSWN----- 320
DB 710 LIKWDELIFPIQSLPLESVLHLLTFLGVLLNOSGSSPDSNKKORKEPALGKVSILTFDFKRL 769
QY 321 -RAYLMAGMEKNLTVMRIRIRIKESGSGOPDNDPILDYGGVDFRLYQLENKSNISGTVR 379
DB 770 FKRLTGC-----TKLLYLW-----TSSHTNSIPGAIIPKSYVMERIVLQVDFPSPADITY 816
QY 380 SGVYVNPBSGKALQLDVYVPLGKIGSGYFOIFOGYGSOLIDVNHGATSGV 432
DB 817 FDIYTSQIDRNIIQDKLETLESIDKG-----KLLDIHRDSSGFL 859

RESULT 5
US-09-170-996-13
; Sequence 13, Application US/09170996
; Patent No. 6291220
; GENERAL INFORMATION:
```

APPLICANT: Williams, Lewis T.
APPLICANT: Moiz, Lisa
APPLICANT: Chen, Yen-Wen
TITLE OF INVENTION: No. 6291220e1 PI 3-Kinase Polypeptides
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/170,996
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/609,049
FILING DATE: 29-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-063700US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 1658 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-170-996-13

Query Match 4.2%; Score 98.5; DB 4; Length 1658;
Best Local Similarity 18.9%; Pred. No. 0.81;
Matches 101; Conservative 78; Mismatches 195; Indels 159; Gaps 23;

QY 5 LSTLTSLSCFAILAIQ--QAKAVPNPAFVDEVRS-----ENDLGQDNE 48
DB 381 LSPVTVQRMGEMASVKYSIEIGLQLPVFTCDVSSIVEIIMQALGMVHDDLQ---- 437
QY 49 LPIDVOS-----ATQSASTDTNPDLDEHE--PELYTTALENKTMLINCALNODIMRLAC 101
DB 438 --VDVGSYILKVGQGEVLQNNHCISHEHIONCRKMDPEIRKIQLLTSLAMQNLARTAE 495
QY 102 VDTLVHGETPAVTKRSTRIDETIMQTIKGRPOVVOYQETTPDIF----- 146
DB 496 DD-----EAP-----VDLKKYILQIEKPYKEVMIRHPVEELDSYHYOVELAQTEN 542
QY 147 -----LMGNEKGLTKRKDAKOLEYAA-----KQFTPLSTSFDDRRNN 183
DB 543 QHRAVDQVIKAVRKICSAIDGVETPSVTEA--VKKILRAVNLNRNSADVTSGSTTRKN 601
QY 184 TPLMSSRPINPMYVLP-----IFMGRPNR---SPMTPSHEARQFTPNNEFRAD 229
DB 602 STKGSIMPENPQVSWDHLTTAIDLLRLHANSSRCSTGCPGSRRIKEAWTATE----- 656
QY 230 LKFOYSVKVKAEDL--WGTOSDLMEFGTQOSHMOIFNGKN--SRF---RHVDYQPELF 282
DB 657 ---QLOFTVYAAHGSSNVSNEKYYLLCSLSH---NGKDLFKIQSKVGTYYKNFEY 709
QY 283 LTVQ-----PV--YSDLPMWDGKVRMIGMGAHVHNSGEASAKLSRNV----- 320
DB 710 LKMWELIIFPIQIOLPLESVLIHLTLFGVLANOSSGSSDSSKOKRGPRALGKVSLLTLD 769

QY 321 -RAYLMAGNEMKNNLVMPRMGRIFKEGSGOPDDNPILDYGYGVDFVLEKNSNI 379
DB 770 KRFELTC-----TKLILY-----TSHTNSIPGAIIPKRSYMERIVLQVDFPSPA 816
QY 380 SGTVRNPRSGGALQDLYVPLGKISGYFQIFQYGGSLDYNHEARFQV 432
DB 817 FDIYTSPOIDRNIIOQDKLETLESIDKG-----KLIDIIHRSSFGL 859

RESULT 6

US-09-091-117-5
Sequence 5, Application US/09091117
Patent No. 6171589
GENERAL INFORMATION:
APPLICANT: The University of Melbourne
TITLE OF INVENTION: Mycoplasma Recombinant Polypeptides and
TITLE OF INVENTION: Vaccines
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: GREENLEE, WINNER and SULLIVAN P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: United States of America
ZIP: 80503
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/091,117
FILING DATE: 12 JUNE 1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU96/00803
FILING DATE: 13-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PN7127
FILING DATE: 13-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: WINNER, Ellen P.
TELECOMMUNICATION INFORMATION:
TELEPHONE: +1 303 499 8080
TELEFAX: +1 303 499 8089
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1024 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Mycoplasma genitalium
US-09-091-117-5

Query Match 3.8%; Score 90.5; DB 4; Length 1024;
Best Local Similarity 20.2%; Pred. No. 2.6;
Matches 86; Conservative 61; Mismatches 172; Indels 107; Gaps 22;

QY 1 MKVS-LSTLTSLSCFAILAIQAKAVPNPAFVDEVSENDLGQDNEPLT-DVQSAHQ 58
DB 1 MKISTTTTCLISLSCAFGTTAI---ALPTVALLKNHQQQNTKQON--PIKIDRFGIN 54
QY 59 SASDTANPLDEHEPELTYTTALENKTMLINCALNODIMRLACYDPLVHGETPAVTKR 118
DB 55 NQAVPTTIL--HQYVEVT--NNKAIYKDAPOKFFL-----AKSALNNKL 98
QY 119 SIRLDETIMQTIKGRPOVVOYQETTPDIFLMGNEKGLTKRKDAKOLEYAKO--FTPLSLSP 177
DB 99 QVEFDFILRT-----GVINLNDLKKIMIDQTLFIPQSF 135
QY 178 DLDNRNTPLMSSRPINPMYVLP--IFMGRPNRNP--NTPSHEARQFTPNNEFRABELFQVS 235

Db 136 DLSANKNLTLNSQSEVSLDLEFIETNFSDKNQPLKLPDGSVYVANE-----SYTSYK 191
Qy 236 VKAKAEDIMGTSDLMFYTQOSHWOIFNGKNSRPRVHDYQPEI-FLTQPYSDL-PM 293
Db 192 ATLQKTKVLYTSHADSHVGSVTAIPTVSLNGKTQNDSEFNPFSKNFPAKNYNNALNPF 251
Qy 294 DGRVIRIGCAVHNS--NGESAK-----LSRSMNRAVYLMAGBWKMLVTMPRIWGRIFK 345
Db 252 EAOQVYVGOGKFLNOKVYNADVDKNDINNHIETQFNVAKITA-----TLGKAFK 300
Qy 346 ---EGSGSQP-----DDNPDLIDY--GYGDV-----RELYOL-- 373
Db 301 QCEHKNQGPLSLKVLKVLGSLNNEFKQLFNVRGLDGFVSDDLQSSQSSNKKTVQQLF 360
Qy 374 EKSNKI 379
Db 361 ENKTTI 366

RESULT 7
US-07-717-331F-2
; Sequence 2, Application US/0717331F
; Patent No. 5484905
; GENERAL INFORMATION:
; APPLICANT: June Nasrallah; Michael Nasrallah; and Joshua
; APPLICANT: Stein
; TITLE OF INVENTION: A Receptor Protein Kinase Gene
; TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Trumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/717,331F
; FILING DATE: June 19th 1991
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)268-1951
; TELEFAX: (203)268-1951
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 857 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-717-331F-2

Query Match 3.8%; Score 89.5; DB 1; Length 857;
Best Local Similarity 21.4%; Pred. No. 2.4; Mismatches 110; Indels 119; Gaps 23;
Matches 76; Conservative 50;

Qy 105 LVHGETPAVYIKTKRSIRLDETIWQITKGRPOVVOYQETPDPIFLMG---NEK---GMLTK 157
Db 24 LIHPALSIYINT-----LSSTESLFTISSNKTIV---SPGSIFFVGFRITNSRYMLGWMYK 75
Qy 158 K-DAKOLEVAAKQFTPLSLSD-----LDNNNTPLWSSRPHNPATVLPITRMCK 205
Db 76 KVSDRYVYVAMNDPLSNAGITLKISGNNVLVDHSHNRPVMTN-----LTRGN 125
Qy 206 PNRSP-----NTPSHEARQFT-----PNEFRAPELKFOYSVKAKAEDLW 245

Db 126 -ERSPVYAEILLANCFVMRSSNNDASEYLMQSRDYDTDLLPBMKGYIKT----- 177
Qy 246 GTDSDLMFEGYTQOSHWOIFNGKNSRPRVHDYQ-----PEIFLTQ---PVYSDLPMDGK 296
Db 178 GLNRFL-----TSMRSSDPSGNGF---SYKLETQSLPEFLYSRNPFPHRRSGPWNG- 226
Qy 297 VRIHGMCAVH-----HSNGESAKLSRSMNRAI-----LMAGBWKMLVTMP--RI 339
Db 227 IIRSGIPEDDKLSYMYNFIENNEVAVTFMTNNSFSRLTILSEGYFQRLATWYPSIRI 286
Qy 340 WGRIFKEGSGSQPDD-----NPDLIDYGYGDVREPLYOLEKNSNIGTVR-YNPRS 389
Db 287 WNRFWSSPVDRQCDTYIMCGP-----YATCDV-----NTSPVCNCIOGFNPRN 329

RESULT 8
US-08-633-879C-2
; Sequence 2, Application US/08633879C
; Patent No. 5928922
; GENERAL INFORMATION:
; APPLICANT: Kivirikko, Kari I.
; APPLICANT: Pihlajaniemi, Taina
; APPLICANT: Heilaakoski, Tarja I.
; APPLICANT: Annunen, Pia P.
; APPLICANT: Nissi, Ritva K.
; APPLICANT: No. 5928922Jelainen, Minna K.
; TITLE OF INVENTION: 2 SUBUNIT OF PROLYL-4-HYDROXYLASE
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING SUCH SUBUNIT AND
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: Fastseq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/633,879C
; FILING DATE: 10-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Abrams, Samuel B.
; REGISTRATION NUMBER: 30,605
; REFERENCE/DOCKET NUMBER: 8389-0041-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 537 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-633-879C-2

Query Match 3.7%; Score 88; DB 2; Length 537;
Best Local Similarity 21.4%; Pred. No. 1.6;
Matches 100; Conservative 52; Mismatches 145; Indels 170; Gaps 26;

TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 503 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-483-322-2

Query Match 3.7%; Score 87.5; DB 1; Length 503;
Best Local Similarity 18.4%; Pred. No. 1.6;
Matches 57; Conservative 31; Mismatches 97; Indels 125; Gaps 14;

QY 40 ENDIGDGNELPIDVQASATDTPANPLDEHEPELYTTALENKTMLNCSALNODIMRL 99
DB 169 QEDIDASNIIDEVSS-----GSTIEKSTPEGYILHTDLP-----SQTGDRD----- 212
QY 100 ACYDTLVHGETPAVYIKRKRSI-----RLDETIWQTIKGRQVYVYQETDPIFLMGNEKG 153
DB 213 ---DAFPGSTLATIATTPWVSAHTKONQERTOWNPIHSNPEVLLQTT----- 258
QY 154 MLTKKDAKOLEYAKOFTPLSLSFDRNNTPL-----WSSRPNPMVYLPFIHMGKPNRS 209
DB 259 -----RMT-----DIDRNSTSAGEWNTQEPQPP----- 283
QY 210 PNTPSHEARQFTPNFPAPELKFOVSVKAAEDLW-----GTSDLMFGYTOQS 259
DB 284 -NNHEYODEEETPH-----ATSTTMADPNSTTEEAATQKEKMF-----EN 322
QY 260 HMOIFNGKN-----SRFRVHDYQPELFIQPYVSDLPW-----DGKVRMIMG 303
DB 323 EMQ---GKNPPTSEDSHVTEGTTASAHNNHPSQRMTOQSOEDVSWDFDPISHPMGOG 379
QY 304 AVHHSNGESA 313
DB 380 HQTESKGHSS 389

RESULT 11
US-08-478-882-2
; Sequence 2, Application US/08478882
; Patent No. 5885575
; GENERAL INFORMATION:
; APPLICANT: HERRLICH, Peter
; APPLICANT: PONTA, Helmut
; APPLICANT: GUNTHER, Ursula
; APPLICANT: MATZKU, Siegfried
; APPLICANT: WENZL, Achim
; TITLE OF INVENTION: VARIANT CD44 SURFACE PROTEINS, DNA
; TITLE OF INVENTION: SEQUENCES CODING THESE, ANTIBODIES AGAINST THESE PROTEINS,
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,882
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/946,497
; FILING DATE: 19921109
; ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16915/145
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 503 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-478-882-2

Query Match 3.7%; Score 87.5; DB 2; Length 503;
Best Local Similarity 18.4%; Pred. No. 1.6;
Matches 57; Conservative 31; Mismatches 97; Indels 125; Gaps 14;

QY 40 ENDIGDGNELPIDVQASATDTPANPLDEHEPELYTTALENKTMLNCSALNODIMRL 99
DB 169 QEDIDASNIIDEVSS-----GSTIEKSTPEGYILHTDLP-----SQTGDRD----- 212
QY 100 ACYDTLVHGETPAVYIKRKRSI-----RLDETIWQTIKGRQVYVYQETDPIFLMGNEKG 153
DB 213 ---DAFPGSTLATIATTPWVSAHTKONQERTOWNPIHSNPEVLLQTT----- 258
QY 154 MLTKKDAKOLEYAKOFTPLSLSFDRNNTPL-----WSSRPNPMVYLPFIHMGKPNRS 209
DB 259 -----RMT-----DIDRNSTSAGEWNTQEPQPP----- 283
QY 210 PNTPSHEARQFTPNFPAPELKFOVSVKAAEDLW-----GTSDLMFGYTOQS 259
DB 284 -NNHEYODEEETPH-----ATSTTMADPNSTTEEAATQKEKMF-----EN 322
QY 260 HMOIFNGKN-----SRFRVHDYQPELFIQPYVSDLPW-----DGKVRMIMG 303
DB 323 EMQ---GKNPPTSEDSHVTEGTTASAHNNHPSQRMTOQSOEDVSWDFDPISHPMGOG 379
QY 304 AVHHSNGESA 313
DB 380 HQTESKGHSS 389

RESULT 12
US-08-961-083-2
; Sequence 2, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders
 REGISTRATION NUMBER: 36,373
 REFERENCE/DOCKET NUMBER: PB340P2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 666 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-961-083-2

Query Match 3.7%; Score 87.5; DB 4; Length 666;
 Best Local Similarity 20.1%; Pred. No. 2.6;
 Matches 88; Conservative 64; Mismatches 188; Indels 97; Gaps 24;

25 KAVNPVAVFVDEVSNDI-GGDNELPIDVQSATOS-----ASTDPANPLDHEPE 74
 1 KIYDKNNOIADIGSERRRVNAQANDIPTLVKAIYSIEDHRRFDHNGITIRILGAFLRN 60
 75 LYTALENKTMILNCSALNODIMRLACYDYLHGETPAVYIKTKRSIRLDETIWQIKGRP 134
 61 LOSNSILOGS-----SLTQOLIKLYFSTSTSDQITIS-----RKAQEAWLAIQLDQ 106
 135 QVVYQET-----IDPIFLMENEKGMILK-----KDAKOLE-----YAAKFTPL 173
 107 KATKOILITLYINKYMSNGTGMQTAQNYGKDLNLSLQALLAGMPQAPNQYDYP 166
 174 S-LSEFDLNNTPMLSSRPHNPMYVLPIMHGKPNRSPNTPSHARQFTPNFEFRPE--- 229
 167 SHPEAQAQDRNLVLT--SEKNGGYI-----SAGEYKAVNTPITDGLQSLKSNAYAYMD 220
 230 --LKFOVVS-VKKAABEDLMGTSDDLMEGYTQ--QSH-WOIFNGKNSRPRVHDYQPEIFL 283
 221 NTLKEVINQVEETGYNLTITGMDVYTNVDOEAKHLMIDYNTDEYVAVPDELQ----- 275
 284 TQPVYSDDLPMDCGVRMIGMAVHNSN-----GESAKLSRSMRAVLAMGMEKKNLTV 336
 276 VASTIYDVS-NGKV-IAQIGARHSSNVSGINQAVETNRDW-----GSTMKPLTDDYA 326
 337 PRI-----WGRIFKESGSGQPDNDPILDY-YGY-GDVRFLYQLENKSNISGTVRYN- 386
 327 PALEYGVYDSTATIYHDEYNYNGTNTPYVNMDRGYFGNITLQYALQGSRRNPVAVETLTK 386
 387 -----PRSGKALQLDY 398
 387 VGLNRAKTFNLGLIDY 403

RESULT 13
 US-08-373-134D-2
 Sequence 2, Application US/08373134D
 Patent No. 5780296

GENERAL INFORMATION:
 APPLICANT: Kmelec, Eric
 APPLICANT: Hollioman, William
 TITLE OF INVENTION: COMPOSITIONS AND METHODS TO PROMOTE
 TITLE OF INVENTION: HOMOLOGOUS RECOMBINATION IN EUKARYOTIC CELLS AND ORGANISMS
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/373,134D
 FILING DATE: January 17, 1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Friedel, Thomas E.
 REGISTRATION NUMBER: 29,258
 REFERENCE/DOCKET NUMBER: 7991-007
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-9741/8864
 TELEX: 66141 PENNITE
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 781 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-373-134D-2

Query Match 3.6%; Score 86; DB 1; Length 781;
 Best Local Similarity 20.7%; Pred. No. 4.9;
 Matches 50; Conservative 37; Mismatches 122; Indels 32; Gaps 9;

39 SENDLGDNELPIDVOS-ATQASSTDPANPLDHEPELYTTALENKTMILNCSALNODIM 97
 533 ARNDPQTSSQAPTSARFSGTGSLPQPLAMDVASQTAFTSGLIASIPLAEAVAREI 592
 98 RLACYDYLHGETPAVYIKTKRSIRLDETIWQI-----KGRPOVYQET-----TTDPI 145
 593 DSAC-----ASNDVPLRTLEARTQLOGT--WSNLINRVFLSKTRARICHRDDQAPACEV 647
 146 FLMGNEKGMILTKDAQOLEYAAQFTPLSLF-FDLDRNNTPLMSSR-----PHNPMYVLP 200
 648 RQNTNORGTASKSLMNTVRAAVVINFPGATMLDVGVDSKSLRQLRFTVTPRAKAVHL-- 705
 201 FMHGKPNRSPNTPSHARQFTPNFEFRAPELKFOVSVKAAEDLMGTSDDLMEGYTQOSH 260
 706 -----NATPSTYHW-AMHATDSTPAEPESQOQRAAERHRAEQADDDLREGALQENH 758
 261 W 261
 759 W 759

RESULT 14
 US-09-114-637-2
 Sequence 2, Application US/09114637
 Patent No. 5945339

GENERAL INFORMATION:
 APPLICANT: Kmelec, Eric
 APPLICANT: Hollioman, William
 TITLE OF INVENTION: COMPOSITIONS AND METHODS TO PROMOTE
 TITLE OF INVENTION: HOMOLOGOUS RECOMBINATION IN EUKARYOTIC CELLS AND ORGANISMS
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/114,637
 FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/373,134
FILING DATE: January 17, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Friedel, Thomas E.
REGISTRATION NUMBER: 29,258
REFERENCE/DOCKET NUMBER: 7991-007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 781 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-114-637-2

Query Match 3.6%; Score 86; DB 2; Length 781;
Best Local Similarity 20.7%; Pred. No. 4.9;
Matches 50; Conservative 37; Mismatches 122; Indels 32; Gaps 9;

QY 39 SENDLDQDNELPIDVOS-ATQASSTDTANPLDHEPELITTALENTMLINCSALNODIM 97
DB 533 ARNDPPTSGAPTSAPSGSGSALPDQPLMDVASQTAFISGLIASIAPTLAEVAGREL 592
QY 98 RLACYDTLVHGETPAVYIKTKRSIRLDETIMQTI-----KGGPOVVOE---TTDPI 145
DB 593 DSAC-----ASNDVPLRLTKLRTAQLGOT-WSNLINRVFLSKTRARICMRDDQAPACEPV 647
QY 146 IFMGNEKMLTKKDAKOLEYAAKQFTPLSL-EDLDENNTPLWSSR---PHNPMVYLP 200
DB 648 RQTNORGTASKSLMTNVRKAAYVINDFGATMLDVGVDKALQLRVLTPRAVHVL- 705
QY 201 FMHGKRNRSNTPSHARQFTNEFRAPELKFOYSKVAKAEDIMGDSLMGTYTQOSH 260
DB 706 -----NAYPTVYM-AMHATADSTPAPESOQOORAAERHPABOEDADDLFGBALQEH 758
QY 261 W 261
DB 759 W 759

RESULT 15
US-08-633-768A-1
Sequence 1, Application US/08633768A
Patent No. 6013504

GENERAL INFORMATION:
APPLICANT: YU, SHUKUN
APPLICANT: BOJSEN, KIRSTEN
APPLICANT: KRAGH, KARSTEN
APPLICANT: BOJKO, MAJA
APPLICANT: NIELSEN, JOHN
APPLICANT: MARCUSSEN, JAN
TITLE OF INVENTION: ALPHA-1,4-GLUCAN LYASE FROM
TITLE OF INVENTION: A FUNGUS INFECTED ALGAE, ITS PURIFICATION
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knudde, Martens, Olsson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/633,768A
FILING DATE: 02-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9321301.5
FILING DATE: 15-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E.
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: DY007.001APC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714-760-0404
TELEFAX: 714-760-9502
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1088 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-633-768A-1

Query Match 3.6%; Score 86; DB 3; Length 1088;
Best Local Similarity 21.6%; Pred. No. 8.4; 138; Indels 84; Gaps 15;
Matches 72; Conservative 40; Mismatches 138; Indels 84; Gaps 15;

QY 106 VHGETPAV-IKTKRSIRLDETIMQTIKGPQVYQETDPIFLMGNEKMLTKKDAKOLE 164
DB 29 IHSAPVAVHATRTKTNLNL-VSMATLSDKOTAFAGSTD-----NDDGI---DYTTYD 76
QY 165 YAAK-QETPLS-----ISFDLDNRNTPPLWSSRPHNPMYVLPDIEM 202
DB 77 YVGWGSFPLSNTEWFAAGSSTPGITDWTAFMNVNEDRIDNP--STVQHPVOVQVTSY 134
QY 203 HGKPNRSNPNTPSHARQFTNEFRAPELKFOYSKVAKAEDIMGDSLMGTYTQOSHQ 262
DB 135 NNSSYRVRFNPDDGIRVTV---RGPIKQOLD-WITFOELSECDPGM----- 178
QY 263 IFNGKNSRPPRVHDYQPEIFLTQPVYSDLPWDGKVRMIGMCAVHHSNGESAKLSRSWNRA 322
DB 179 TFTSEGLTFETKDLSTVYIIGNFRTVTRKSDGKVIN-----ENDEVGTAS----- 224
QY 323 YLMAGMEKMLTYVPRIMGRIFKEGSGSPDPNDLIDYGYGVDVFLYQ---LENKSNI 379
DB 225 ---SGNKCRLMEFVDRLYGNAIASVKNKFNNDKAVKQEGFGAGEVCKYODTYILERTGI 281
QY 380 SGT-----VRYN-----PRSGKALQDLDDVYPL 402
DB 282 AMTNYNDNLNMQMDLRPHHDGALNPDYIIPM 315

Search completed: November 30, 2001, 14:17:48
Job time: 225 sec

...

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:24:38 ; Search time 358.98 Seconds
(Without alignments) 341.870 Million cell updates/sec

Title: US-09-787-083-6

Perfect score: 2360
Sequence: 1 MKVSLSTLTLSLSCFALIA.....YNHQATSGVGLMDNMGL 442

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

Pending Patents AA Main: *
1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep: *
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep: *
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep: *
4: /cgn2_6/ptodata/2/paa/US080_COMB.pep: *
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep: *
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep: *
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep: *
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep: *
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep: *
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep: *
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep: *
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep: *
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep: *
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep: *
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep: *
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep: *
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep: *
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep: *
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep: *
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep: *
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep: *
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep: *
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep: *
24: /cgn2_6/ptodata/2/paa/US60_COMB.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2360	100.0	442	21	US-09-787-083-6
2	2352	99.7	442	21	US-09-787-083-2
3	2348	99.5	442	21	US-09-787-083-8
4	2330	98.7	442	21	US-09-787-083-4
5	1810	76.7	346	19	US-09-540-236-3202
6	1785	75.6	375	24	US-60-128-476-3019
7	815	34.5	374	21	US-09-787-084-4
8	812	34.4	375	21	US-09-787-084-2
9	552	23.4	109	19	US-09-540-236-3197

10	552	23.4	132	24	US-60-128-476-3016	Sequence 3016, Ap
11	548.5	23.2	388	17	US-09-328-352-7532	Sequence 7532, Ap
12	372.5	15.8	293	18	US-09-543-681A-4390	Sequence 4390, Ap
13	371	15.7	297	18	US-09-489-039A-8021	Sequence 8021, Ap
14	357	15.1	263	21	US-09-758-460-912	Sequence 912, Ap
15	354	15.0	307	16	US-09-252-691-7841	Sequence 7841, Ap
16	354	15.0	307	16	US-09-252-691C-7841	Sequence 7841, Ap
17	352.5	14.9	282	24	US-60-215-161-7051	Sequence 7051, Ap
18	263	11.1	749	1	PCT-US01-08631-55161	Sequence 55161, A
19	246.5	10.4	355	1	PCT-US98-06371-1342	Sequence 1342, Ap
20	246.5	10.4	355	13	US-08-902-615A-608	Sequence 608, Ap
21	239.5	10.1	356	11	PCT-US97-19575-90	Sequence 90, Ap
22	239.5	10.1	356	11	US-08-759-739-472	Sequence 472, Ap
23	239.5	10.1	356	13	US-08-993-002A-5205	Sequence 5205, Ap
24	173.5	7.4	591	1	PCT-US01-08631-55155	Sequence 55155, A
25	156.5	6.6	251	11	US-08-759-739-504	Sequence 504, Ap
26	156.5	6.6	251	13	US-08-993-002A-5203	Sequence 5203, Ap
27	156.5	6.6	253	11	US-08-759-739-381	Sequence 381, Ap
28	156.5	6.6	253	13	US-08-993-002A-5204	Sequence 5204, Ap
29	127.5	5.4	134	10	US-08-625-811-1275	Sequence 1275, Ap
30	114	4.8	1686	17	US-09-355-160A-2	Sequence 2, Ap
31	114	4.8	1686	17	US-09-355-160B-2	Sequence 2, Ap
32	107	4.5	157	8	US-08-487-032A-837	Sequence 837, Ap
33	107	4.5	157	9	US-08-561-469A-837	Sequence 837, Ap
34	107	4.5	157	11	US-08-759-739-327	Sequence 327, Ap
35	107	4.5	157	13	US-08-993-002A-5202	Sequence 5202, Ap
36	104.5	4.4	537	21	US-09-738-626-4754	Sequence 4754, Ap
37	104.5	4.4	546	20	US-09-603-124B-410	Sequence 410, Ap
38	104.5	4.4	1030	24	US-60-324-109-30815	Sequence 30815, A
39	104	4.4	798	16	PCT-US99-28332-5	Sequence 5, Ap
40	104	4.4	798	16	US-09-203-453-5	Sequence 5, Ap
41	104	4.4	798	21	US-09-712-032-5	Sequence 5, Ap
42	104	4.4	798	23	US-09-948-947-86	Sequence 86, Ap
43	104	4.4	1414	24	US-60-314-050-7662	Sequence 7662, Ap
44	101.5	4.3	467	1	PCT-US01-04098A-1004	Sequence 1004, Ap
45	101.5	4.3	467	14	US-09-062-009-6	Sequence 6, Ap

ALIGNMENTS

RESULT 1
US-09-787-083-6
; Sequence 6, Application US/09787083
; GENERAL INFORMATION:
; APPLICANT: Rueille, Jean-Louis
; TITLE OF INVENTION: Novel Compounds
; FILE REFERENCE: BM45332
; CURRENT APPLICATION NUMBER: US/09/787,083
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: PCT/EP99/06781
; PRIOR FILING DATE: 1999-09-14
; PRIOR APPLICATION NUMBER: GB 9820002.5
; PRIOR FILING DATE: 1998-09-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-09-787-083-6

Query Match 100.0%; Score 2360; DB 21; Length 442;
Best Local Similarity 100.0%; Pred. No. 4.2e-235;
Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MKVSLSTLTLSLSCFALIAQAKAVNPVAFVDEVSNDLGDNELPTDVOSATQSA 60
Db 1 MKVSLSTLTLSLSCFALIAQAKAVNPVAFVDEVSNDLGDNELPTDVOSATQSA 60
OY 61 STDPAHPDEHPPELYTALNKMTLINCALNDIMLACYDPLVHGEPVAVIKTKRSI 120
|||||

Db 61 STDTANPLDEHEBELYTTALENKTMLINCSALNODIMRLACYDTLVHGETPAVYIKTKRSI 120
QY 121 RLDETIMQTIKGRPOVYVOETTDPIFLMGNEKMLTKKDAKOLEYAAKQFTPLSLSFDD 180
Db 121 RLDETIMQTIKGRPOVYVOETTDPIFLMGNEKMLTKKDAKOLEYAAKQFTPLSLSFDD 180
QY 181 RNNTPPLMSRPHNPMYVLPFIEMHGKPNRSPNTPSHAKQFTPREFAPBLKFOVSVKYVA 240
Db 181 RNNTPPLMSRPHNPMYVLPFIEMHGKPNRSPNTPSHAKQFTPREFAPBLKFOVSVKYVA 240
QY 241 AEDLMGTDSDLMEGYTQOSHMOIFNGKNSRPFVHDYQPEIFLTQPVYSDLPMDGKYRMI 300
Db 241 AEDLMGTDSDLMEGYTQOSHMOIFNGKNSRPFVHDYQPEIFLTQPVYSDLPMDGKYRMI 300
QY 301 GMAVHNSGESAKLSRSNNRAYLMAEMKMLTVMRIMGRIFKSGSSQPDNDPDIID 360
Db 301 GMAVHNSGESAKLSRSNNRAYLMAEMKMLTVMRIMGRIFKSGSSQPDNDPDIID 360
QY 361 YYGVDVRFYLOLEKNSNIGTVRYNPRSGKALQDYYVPLGKISGYFOIGYGQSL 420
Db 361 YYGVDVRFYLOLEKNSNIGTVRYNPRSGKALQDYYVPLGKISGYFOIGYGQSL 420
QY 421 IDYNHEATSEFGVGLMLNDMMGL 442
Db 421 IDYNHEATSEFGVGLMLNDMMGL 442

RESULT 2

US-09-787-083-2
; Sequence 2, Application us/09787083
; GENERAL INFORMATION:
; APPLICANT: Ruelle, Jean-Louis
; TITLE OF INVENTION: Novel Compounds
; FILE REFERENCE: BM45332
; CURRENT APPLICATION NUMBER: US/09/787,083
; PRIORITY FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: PCT/EP99/06781
; PRIOR FILING DATE: 1999-09-14
; PRIOR APPLICATION NUMBER: GB 9820002.5
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 2
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-09-787-083-2

Query Match 99.7%; Score 2352; DB 21; Length 442;
Best Local Similarity 99.3%; Pred. No. 2,8e-234;

Matches 439; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVSLSTLTLISLSCFAIILAIQAKAVPNPVAFVDEVSENDLGONELPIDVQATQSA 60
Db 1 MKVSLSTLTLISLSCFAIILAIQAKAVPNPVAFVDEVSENDLGONELPIDVQATQSA 60
QY 61 STDTANPLDEHEBELYTTALENKTMLINCSALNODIMRLACYDTLVHGETPAVYIKTKRSI 120
Db 61 STDTANPLDEHEBELYTTALENKTMLINCSALNODIMRLACYDTLVHGETPAVYIKTKRSI 120
QY 121 RLDETIMQTIKGRPOVYVOETTDPIFLMGNEKMLTKKDAKOLEYAAKQFTPLSLSFDD 180
Db 121 RLDETIMQTIKGRPOVYVOETTDPIFLMGNEKMLTKKDAKOLEYAAKQFTPLSLSFDD 180
QY 181 RNNTPPLMSRPHNPMYVLPFIEMHGKPNRSPNTPSHAKQFTPREFAPBLKFOVSVKYVA 240
Db 181 RNNTPPLMSRPHNPMYVLPFIEMHGKPNRSPNTPSHAKQFTPREFAPBLKFOVSVKYVA 240
QY 241 AEDLMGTDSDLMEGYTQOSHMOIFNGKNSRPFVHDYQPEIFLTQPVYSDLPMDGKYRMI 300
Db 241 AEDLMGTDSDLMEGYTQOSHMOIFNGKNSRPFVHDYQPEIFLTQPVYSDLPMDGKYRMI 300

QY 301 GMAVHNSGESAKLSRSNNRAYLMAEMKMLTVMRIMGRIFKSGSSQPDNDPDIID 360
Db 301 GMAVHNSGESAKLSRSNNRAYLMAEMKMLTVMRIMGRIFKSGSSQPDNDPDIID 360
QY 361 YYGVDVRFYLOLEKNSNIGTVRYNPRSGKALQDYYVPLGKISGYFOIGYGQSL 420
Db 361 YYGVDVRFYLOLEKNSNIGTVRYNPRSGKALQDYYVPLGKISGYFOIGYGQSL 420
QY 421 IDYNHEATSEFGVGLMLNDMMGL 442
Db 421 IDYNHEATSEFGVGLMLNDMMGL 442

RESULT 3

US-09-787-083-8
; Sequence 8, Application us/09787083
; GENERAL INFORMATION:
; APPLICANT: Ruelle, Jean-Louis
; TITLE OF INVENTION: Novel Compounds
; FILE REFERENCE: BM45332
; CURRENT APPLICATION NUMBER: US/09/787,083
; PRIORITY FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: PCT/EP99/06781
; PRIOR FILING DATE: 1999-09-14
; PRIOR APPLICATION NUMBER: GB 9820002.5
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 8
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-09-787-083-8

Query Match 99.5%; Score 2348; DB 21; Length 442;
Best Local Similarity 99.3%; Pred. No. 7,3e-234;

Matches 439; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKVSLSTLTLISLSCFAIILAIQAKAVPNPVAFVDEVSENDLGONELPIDVQATQSA 60
Db 1 MKVSLSTLTLISLSCFAIILAIQAKAVPNPVAFVDEVSENDLGONELPIDVQATQSA 60
QY 61 STDTANPLDEHEBELYTTALENKTMLINCSALNODIMRLACYDTLVHGETPAVYIKTKRSI 120
Db 61 STDTANPLDEHEBELYTTALENKTMLINCSALNODIMRLACYDTLVHGETPAVYIKTKRSI 120
QY 121 RLDETIMQTIKGRPOVYVOETTDPIFLMGNEKMLTKKDAKOLEYAAKQFTPLSLSFDD 180
Db 121 RLDETIMQTIKGRPOVYVOETTDPIFLMGNEKMLTKKDAKOLEYAAKQFTPLSLSFDD 180
QY 181 RNNTPPLMSRPHNPMYVLPFIEMHGKPNRSPNTPSHAKQFTPREFAPBLKFOVSVKYVA 240
Db 181 RNNTPPLMSRPHNPMYVLPFIEMHGKPNRSPNTPSHAKQFTPREFAPBLKFOVSVKYVA 240
QY 241 AEDLMGTDSDLMEGYTQOSHMOIFNGKNSRPFVHDYQPEIFLTQPVYSDLPMDGKYRMI 300
Db 241 AEDLMGTDSDLMEGYTQOSHMOIFNGKNSRPFVHDYQPEIFLTQPVYSDLPMDGKYRMI 300
QY 301 GMAVHNSGESAKLSRSNNRAYLMAEMKMLTVMRIMGRIFKSGSSQPDNDPDIID 360
Db 301 GMAVHNSGESAKLSRSNNRAYLMAEMKMLTVMRIMGRIFKSGSSQPDNDPDIID 360
QY 361 YYGVDVRFYLOLEKNSNIGTVRYNPRSGKALQDYYVPLGKISGYFOIGYGQSL 420
Db 361 YYGVDVRFYLOLEKNSNIGTVRYNPRSGKALQDYYVPLGKISGYFOIGYGQSL 420
QY 421 IDYNHEATSEFGVGLMLNDMMGL 442
Db 421 IDYNHEATSEFGVGLMLNDMMGL 442

RESULT 4

```
US-09-787-083-4
; Sequence 4, Application US/09787083
; GENERAL INFORMATION:
; APPLICANT: Rueille, Jean-Louis
; TITLE OF INVENTION: Novel Compounds
; FILE REFERENCE: BM45332
; CURRENT APPLICATION NUMBER: US/09/787, 083
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: PCT/EP99/06781
; PRIOR FILING DATE: 1999-09-14
; PRIOR APPLICATION NUMBER: GB 9820002.5
; PRIOR FILING DATE: 1998-09-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-09-787-083-4
```

```
Query Match          98.7%; Score 2330; DB 21; Length 442;
Best Local Similarity 98.9%; Pred. No. 5.4e-232;
Matches 437; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 MKVSLSTLTLSILSCFAILAIQAKAVPNVAFVDEVRENDLGODNELPIDVOSATQSA 60
    |||||||
DB 1 MKVSLSTLTLSILSCFAILAIQAKAVPNVAFVDEVRENDLGODNELPIDVOSATQSA 60
QY 61 STDANPLDEHEPELYTTALENKMTMLINCSALNODIMRLACYDTLVHGEPYAVIKTKRSI 120
    |||||||
DB 61 STDANPLDEHEPELYTTALENKMTMLINCSALNODIMRLACYDTLVHGEPYAVIKTKRSI 120
QY 121 RLDETIWQTIKGRPOVYVYQETTDPIFLMGNEKMLTKKDAKOLEYAAKQFTPLSLSPDL 180
    |||||||
DB 121 RLDETIWQTIKGRPOVYVYQETTDPIFLMGNEKMLTKKDAKOLEYAAKQFTPLSLSPDL 180
QY 181 RNNTPPLMSSRPNNPMYVLPFPMHGKPNRSPPTPSHEAKQFTPNFRAPELKFQVSVKYA 240
    |||||||
DB 181 RNNTPPLMSSRPNNPMYVLPFPMHGKPNRSPPTPSHEAKQFTPNFRAPELKFQVSVKYA 240
QY 241 AEDLMGTDSDLMFGYTOOSHMOIFNGKNSRPFVHDYQPELFTLPQVYSDLPMDGKYRMI 300
    |||||||
DB 241 AEDLMGTDSDLMFGYTOOSHMOIFNGKNSRPFVHDYQPELFTLPQVYSDLPMDGKYRMI 300
QY 301 GMGAVHNSGESAKLSRSNNRAYLMAEMKNTLVMPRINGRIKKEGSGQPDNDPILD 360
    |||||||
DB 301 GMGAVHNSGESAKLSRSNNRAYLMAEMKNTLVMPRINGRIKKEGSGQPDNDPILD 360
QY 361 YYGVDVRFYLOLEKNSNIGTVARYNPRSGKALQLDYVYPLGKIGISGYFOIGYQOSL 420
    |||||||
DB 361 YYGVDVRFYLOLEKNSNIGTVARYNPRSGKALQLDYVYPLGKIGISGYFOIGYQOSL 420
QY 421 IDYNHEATSFVGGLMLNDMMGL 442
    |||||||
DB 421 IDYNHEATSFVGGLMLNDMMGL 442
```

```
RESULT 5
US-09-540-236-3202
; Sequence 3202, Application US/09540236
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: 2709, 2005-001
; CURRENT APPLICATION NUMBER: US/09/540, 236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3202
; LENGTH: 346
; TYPE: PRT
; ORGANISM: M. catarrhalis
```

```
US-09-540-236-3202
```

```
Query Match          76.7%; Score 1810; DB 19; Length 346;
Best Local Similarity 99.4%; Pred. No. 3.2e-178;
Matches 339; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MKVSLSTLTLSILSCFAILAIQAKAVPNVAFVDEVRENDLGODNELPIDVOSATQSA 60
    |||||||
DB 3 MKVSLSTLTLSILSCFAILAIQAKAVPNVAFVDEVRENDLGODNELPIDVOSATQSA 62
QY 61 STDANPLDEHEPELYTTALENKMTMLINCSALNODIMRLACYDTLVHGEPYAVIKTKRSI 120
    |||||||
DB 63 STDANPLDEHEPELYTTALENKMTMLINCSALNODIMRLACYDTLVHGEPYAVIKTKRSI 122
QY 121 RLDETIWQTIKGRPOVYVYQETTDPIFLMGNEKMLTKKDAKOLEYAAKQFTPLSLSPDL 180
    |||||||
DB 123 RLDETIWQTIKGRPOVYVYQETTDPIFLMGNEKMLTKKDAKOLEYAAKQFTPLSLSPDL 182
QY 181 RNNTPPLMSSRPNNPMYVLPFPMHGKPNRSPPTPSHEAKQFTPNFRAPELKFQVSVKYA 240
    |||||||
DB 183 RNNTPPLMSSRPNNPMYVLPFPMHGKPNRSPPTPSHEAKQFTPNFRAPELKFQVSVKYA 242
QY 241 AEDLMGTDSDLMFGYTOOSHMOIFNGKNSRPFVHDYQPELFTLPQVYSDLPMDGKYRMI 300
    |||||||
DB 243 AEDLMGTDSDLMFGYTOOSHMOIFNGKNSRPFVHDYQPELFTLPQVYSDLPMDGKYRMI 302
QY 301 GMGAVHNSGESAKLSRSNNRAYLMAEMKNTLVMPRING 341
    |||||||
DB 303 GMGAVHNSGESAKLSRSNNRAYLMAEMKNTLVMPRING 343
```

```
RESULT 6
US-60-128-476-3019
; Sequence 3019, Application US/60128476
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA
; FILE REFERENCE: 107196, 139
; CURRENT APPLICATION NUMBER: US/60/128, 476
; CURRENT FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 5002
; SEQ ID NO 3019
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-60-128-476-3019
```

```
Query Match          75.6%; Score 1785; DB 24; Length 375;
Best Local Similarity 98.0%; Pred. No. 1.5e-175;
Matches 336; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 MKVSLSTLTLSILSCFAILAIQAKAVPNVAFVDEVRENDLGODNELPIDVOSATQSA 60
    |||||||
DB 3 MKVSLSTLTLSILSCFAILAIQAKAVPNVAFVDEVRENDLGODNELPIDVOSATQSA 62
QY 61 STDANPLDEHEPELYTTALENKMTMLINCSALNODIMRLACYDTLVHGEPYAVIKTKRSI 120
    |||||||
DB 63 STDANPLDEHEPELYTTALENKMTMLINCSALNODIMRLACYDTLVHGEPYAVIKTKRSI 122
QY 121 RLDETIWQTIKGRPOVYVYQETTDPIFLMGNEKMLTKKDAKOLEYAAKQFTPLSLSPDL 180
    |||||||
DB 123 RLDETIWQTIKGRPOVYVYQETTDPIFLMGNEKMLTKKDAKOLEYAAKQFTPLSLSPDL 182
QY 181 RNNTPPLMSSRPNNPMYVLPFPMHGKPNRSPPTPSHEAKQFTPNFRAPELKFQVSVKYA 240
    |||||||
DB 183 RNNTPPLMSSRPNNPMYVLPFPMHGKPNRSPPTPSHEAKQFTPNFRAPELKFQVSVKYA 242
QY 241 AEDLMGTDSDLMFGYTOOSHMOIFNGKNSRPFVHDYQPELFTLPQVYSDLPMDGKYRMI 300
    |||||||
DB 243 AEDLMGTDSDLMFGYTOOSHMOIFNGKNSRPFVHDYQPELFTLPQVYSDLPMDGKYRMI 302
```

```
QY      301 GMGAVHSHNGESAKLSRSWNRAYLMAGMEWKNTVMPRIWGRI   343
        |||||  
Db      303 GMGAVHSHNGESAKLSRSWNRAYLMAGMEWKNTVMPTHLGGV   345
```

RESULT 7

```

US-09-787-084-4
: Sequence 4, Application US/09787084
: GENERAL INFORMATION:
: APPLICANT: Ruelle, Jean-Louis
: TITLE OF INVENTION: polynucleotides and polypeptides BASB033
: TITLE OF INVENTION: from Neisseria Meningitidis and Their Uses
: FILE REFERENCE: BM45331
: CURRENT APPLICATION NUMBER: US/09/787, 084
: CURRENT FILING DATE: 2001-03-14
: PRIOR APPLICATION NUMBER: PCT/EP99/06718
: PRIOR FILING DATE: 1999-09-09
: PRIOR APPLICATION NUMBER: GB 9820003.3
: PRIOR FILING DATE: 1998-09-14
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: fastseq for Windows Version 4.0
: SEQ ID NO. 4
: LENGTH: 374
: TYPE: PRT
: ORGANISM: Neisseria meningitidis
: US-09-787-084-4

```

Query Match	34.5%	Score 815	DB 21	Length 374
Best Local Similarity	44.7%	Pred. No. 6.3e-75		
Matches 163	Conservative 67	Mismatches 111	Indels 24	Gaps 8

QY	87	INCSALMODIRRLACYPDVLHNGETPAVI-----KTKRSIRLDETIMQTI-KGRPOVYOE	140
Db	25	LQCAALIDNVTIRLACIOYBIRFAOLPSSAGOGESKAVLWTEVRSLSDKGEAVIYEK	84
QY	141	TTDPITFLMGNEKMLTKKDAKOLEYAAKQFTPLSLSPDLDRNN-TPLMSSRPNNMYPVL	199
Db	85	GGDAL-----PADSAGETADITYTLPISLWYDLDKNDRLGLLVGREHNNPYMLP	131
QY	200	IFMIGKPNRSPNNTSHEAR-QFTTNERAPAEIKQVSVKKKAEDLMGTSDLMFGYTQO	256
Db	132	LMWNNSPYAGSPSTRGTTOVEKQOQKRAETKLOVFSKSIADLEDKTADLMFGYTOR	191
QY	259	SHMOIFN-GKRSRPFVHDYOPETFLQOPAYSDLPMPGKVRMIGMGAVHNSNESAKLSR	317
Db	192	SDMOITINOGKRSAPFRKNTDYKPELFEFLQPVKADLPFGGRULRMGAGVHVHNSQOSPRESR	251
QY	318	SMNRAYVLAMGEMENLTVMPRIWGRIFKEGSGSQPDNDPILDYGGVADRFYLOLEKNS	377
Db	252	SMNITLYAMAGEMCKLTVIPRVWVRALDQ--SGDK-NONPRIDADYMGIGVDKQLQYRLNDRQ	309
QY	378	NISGTVARYNPRSGKALQLDVYVPLGKISGYPQIIFQGYQOSLIDYVNEHETSFGVGIMLN	437
Db	310	NWYSVLNRVNPXTGYATEAATYPIFKKLGKGVNAGVFGYGSGLIDYVNHKNGIGIGLMFN	369
QY	438	DMGGL 442	
Db	370	DLDGI 374	

```

1  RESULT      8
2  US-09-787-084-2
3  Sequence 2, Application US/09767084
4  GENERAL INFORMATION:
5  APPLICANT: Ruelle, Jean-Louis
6  TITLE OF INVENTION: Polynucleotides and Polypeptides BASB033
7  TITLE OF INVENTION: from Neisseria Meningitidis and Their Uses
8  FILE REFERENCE: Bm45331
9  CURRENT APPLICATION NUMBER: US/09/787,084
10 CURRENT FILING DATE: 2001-03-14
11 PRIOR APPLICATION NUMBER: PCT/EP99/06718

```

```

: PRIOR FILING DATE: 1999-09-09
: PRIOR APPLICATION NUMBER: GB 9820003.3
: PRIOR FILING DATE: 1998-09-14
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO: 1

```

Query Match	34.48	Score 812	DB 21	Length 375
Best Local Similarity	44.48	Pred. No. 1.3e-74		
Matches 162	Conservative 68	Mismatches 111	Indels 24	Gaps 8

[illegible]

```

RESULT      9
US-09-540-236-3197
; Sequence 3197, Application US/09540236
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CAPSI
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540.236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3197
; LENGTH: 109
; TYPE: PR1
; ORGANISM: M.cattarrhalis
; US-09-540-236-3197

```

	Query Match	23.4%	Score 552;	DB 19;	Length 109;
	Best Local Similarity	100.0%;	Pred. NO. 1.7e-46;		
	Matches 102; Conservative	0;	Mismatches	0;	Gaps 0
QY	341 GRIFKSGSSGQDDDPDILIDYGYCDVRFYLQLEKNKSINSGTVRYNPRSGKALQLDYVV	400			
Dd	8 GRIFKSGSSGQDDDPDILIDYGYCDVRFYLQLEKNKSINSGTVRYNPRSGKALQLDYVV	67			
QY	401 PLKGISGYSFQIIFGGYGOSLIDYNHEATSFVGVLMDMMGL	442			

Db 68 PLKGISGYFQIFGGYQSLIDYNHEATSFVGGLMLNDWMGL 109

```

RESULT 10
US-60-128-476-3016
: Sequence 3016, Application US/60128476
: GENERAL INFORMATION:
: APPLICANT: Gary L. Breton et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA
: FILE REFERENCE: 107196.139
: CURRENT APPLICATION NUMBER: US/60/128,476
: CURRENT FILING DATE: 1999-04-09
: NUMBER OF SEQ ID NOS: 5002
: SEQ ID NO 3016
: LENGTH: 132
: TYPE: PRN
: ORGANISM: Moraxella catarrhalis
US-60-128-476-3016

```

Query Match	23.4%	Score	552	DB	24	Length	132
Best Local Similarity	100.0%	Pred. No.	2	3e-48			
Matches	102	Conservative	0	Mismatches	0	Indels	0

QY	DB	QY	DB
341	91	31	401
GRIRKESGSGSPDDNPILDYGGVDRFLQLEKNKNSISTVAYNRRSGALQILDYV	PLGKGISGTFQIFQGGSLIDYNNHEATSRVGLMLDMMGL	GRIRKESGSGPDDNPILDYGGVDRFLQLEKNKNSISTVAYNRRSGALQILDYV	PLGKGISGTFQIFQGGSLIDYNNHEATSRVGLMLDMMGL
405	132	90	442
GRIRKESGSGPDDNPILDYGGVDRFLQLEKNKNSISTVAYNRRSGALQILDYV	PLGKGISGTFQIFQGGSLIDYNNHEATSRVGLMLDMMGL	GRIRKESGSGPDDNPILDYGGVDRFLQLEKNKNSISTVAYNRRSGALQILDYV	PLGKGISGTFQIFQGGSLIDYNNHEATSRVGLMLDMMGL

```

RESULT 11
US-09-328-352-7532
: Sequence 7532, Application US/09328352
: GENERAL INFORMATION:
: APPLICANT: Gary L. Breton et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
: TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: GTC99-03PA
: CURRENT APPLICATION NUMBER: US/09/328,352
: CURRENT FILING DATE: 1999-06-04
: NUMBER OF SEQ ID NOS: 8252
: SEQ ID NO 7532
:
: LENGTH: 388
: TYPE: PRT
: ORGANISM: Acinetobacter baumannii
: US-09-328-352-7532

```

Query Match	23.2%	Score 548.5	DB 17	Length 388
Best Local Similarity	34.4%	Pred. No. 3e-47		
Matches 126	Conservative 70	Mismatches 129	Indels 41	Gaps 10

```

QY      89  CSALNODIRLACIDYDLVNGEPAVIKT-----KRSILDEFTIMOTIKGPVOWIQTDP  144
Db      48  CVALASMDRLACDYAVFRPSALPYYQAAVYPEYVKIKDRPVYQPEFFKEKVV--DKSN  105

QY      145  IFLMGNEKGMLTKDKAQLEIYAAKOFTPLSLSPD-----LDRNNTPLMSSRPHNPMVLP  199
Db      106  IKVIVK-----APTLEPTTSLDDQWELSEKSKLGIVNIRAYOPVYLLP  149

QY      200  IFMHGKPNRSPMTPSHEARQFTTNEFRAPELKFOVSVKVAEELMGSDSLMGCTQGS  259
Db      150  VFWTSDKNFEFPSPNNNTVTTEAQNITAKSTESKFOISLKTAMENITFGNNGGLMWGYTSS  209

QY      260  HMOIFNGKRSRFEPRVADYOPE---IFLTQPVYSDLPMDGKVRMIMGAVHHNSGESAFLS  316
Db      210  RMQGTNAEESRRPREFTNVEPEASLMFRTN--YELTLGIDG--RLLGVTPLNHOSNGSDPLS  265

QY      317  RSMNRATYLMAGBEMKULVMPRIWGRIFKEGSGSQPDNDPDIIDYGYGADVRFYOLENK  376

```

```
Db      266 RSMNRVIFNVGLERGNFALMLRPWIRLEED---SKDDNNPDMEDYIGRGDLTAFFYKWK-Q 321
```

Qy 377 SNIGSTVRVPRSG---KGAQLDLYVPLGKGISGYFQIQGCGSLIDYHNHATSGVG 433

Db 322 NDFSLMLRHSLKGGDDSHGAVQFDMAFPISGKLRGHHFQLENGYGESLIDYHNHRTAVAGLG 380

QY	434	LMLNDW	435
	:	:	:
Db	382	VSLMNV	387

```

RESULT 12
US-09-543-681A-4390
: Sequence 4390, Application US/09543681A
: GENERAL INFORMATION:
: APPLICANT: GARY BRETTON
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
: TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 27/09.1002-001
: CURRENT APPLICATION NUMBER: US/09/543.681A
: PRIOR FILING DATE: 2000-04-05
: PRIOR APPLICATION NUMBER: US 60/128,706
: PRIOR FILING DATE: 1999-04-09
: NUMBER OF SEQ ID NOS: 8344
: SEQ ID NO 4390
: LENGTH: 293
: TYPE: PRT
: ORGANISM: Proteus mirabilis
: US-09-543-681A-4390

```

Query Match	15.8%	Score 372.5;	DB 19;	Length 293;
Best Local Similarity	39.6%	Pred. No. 3.4e-29;		
Matches 90;	Conservative 35;	Mismatches 87;	Indels 15;	Gaps 6

```

OY 215 HEAROFTEPNERAEKLTQVSVKTKAAEDDLGTSOTDMEFGTOOSHQJFNGKSNRPFRV 274
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 81 HKAKK-----DEVKQJISLAPFLRMGLNGIGENSVLAASTQTSNMWQJLSNKKKSAFPRE 133
OY 275 HDYOPEIFLTPV--YSDDLPMWDGKVRYMIGMGAHVHNSGEASAKLSRNNRAYLMAEMENKL 333
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 133 TNEPEOLFEMATDYOQFAGW--TLREIETGPNHESNGAEPTSRSMRYVRAFMQAQGNF 190
OY 334 TVMPEIRMTKEGSGSPDDNPOLDLYGCGDYVRLQLEKNKSNIGYVRYNPRSGKA 399
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 191 QLDLKPWR--RNE--SAKNDNPNPDIRMAGYRLKAGRI--GESTFLTLGYTNMWSGYGA 246
OY 394 LOLDVVPYLGKIGSYFQIFQGYOSLIDYHNEATSGVGLMLNDMM 440
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 247 AELGMSYPTKHNRPYTOLESGYESMDYNEPROTRLGVGMALNDML 293

```

```

RESULT 13
US-09-489-039A-8021
: Sequence 8021, Application US/09489039A
: GENERAL INFORMATION:
: APPLICANT: Garry Bretton et. al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
: TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 2709.2004001
: CURRENT APPLICATION NUMBER: US/09/489,039A
: CURRENT FILING DATE: 2000-01-27
: PRIOR APPLICATION NUMBER: US 60/117,747
: PRIOR FILING DATE: 1999-01-29
: NUMBER OF SEQ ID NOS: 14342
: SEQ ID NO 8021
:
: LENGTH: 297
:
: TYPE: PAT
:
: ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8021

```

[illegible]

```

14 RESULT
15 US-09-758-460-912
16 : Sequence 912. Application US/09758460
17 : GENERAL INFORMATION:
18 : APPLICANT: Rosen et al.
19 : TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
20 : FILE REFERENCE: PMO10
21 : CURRENT APPLICATION NUMBER: US/09/758,460
22 : CURRENT FILING DATE: 2001-01-11
23 : PRIOR APPLICATION NUMBER: 60/179,065
24 : PRIOR FILING DATE: 2000-01-31
25 : PRIOR APPLICATION NUMBER: 60/180,628
26 : PRIOR FILING DATE: 2000-02-04
27 : NUMBER OF SEQ ID NOS: 940
28 : SOFTWARE: PatentIn Ver. 2.0
29 : SEQ ID NO 912
30 : LENGTH: 263
31 : TYPE: PRT
32 : ORGANISM: Homo sapiens
33 : FEATURE:
34 : NAME/KEY: SITE
35 : LOCATION: (2)
36 : OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
37 : US-09-758-460-912

```

	Query Match	15.1%	Score 357;	DB 21	Length 263;	
	Best Local Similarity	39.3%	Pred. No. 1,2e-27;			
	Matches	84;	Conservative	29;	Mismatches 91;	Indels 10; Gaps 4
OY	226	RAPELKFOVSVKVAKAEDMGTDSDLMEGYTQOOSHWFNGKNSRPFVHDYOPELTELQ	285			
		: :: :::	:	:	:	:
Db	57	RKDECFPLDLSFLPMLRGILGPNISLGASYTKRSWMOLOSSESSPFRETNPEOLFGE	116			
OY	286	PV-YEYDLPMDDGVRRMIGAGVHNSGESAKLSRSNNRAVLAMGMEEKNLTVPRIMGRFE	344			
Db	117	ATDYRFACH--TLRVEMKYNHDSNGRSDPTSRSMRLYLTRLAENGNALEVKKPYVV-	173			
OY	345	KEGSGSQPDNDNPDLLIDYYGGDVREFLYOLENKSNTISGTVRNPRSRCGALQLDYVPFLK	404			
Db	174	-----GMTDDNDNPDIETRYKMGYOLKIGYHL-GDAVLSAKQYWMNNGYGGAELGLSVPTRK	227			
OY	405	GISCYGFQFQGGOSLDIYNHEATSFVGVMAMD	438			
Db	228	HRLRTQVYSGGESLIDNFNQTRAGGVAMLND	261			

RESULT 15
US-09-252-691-7841
; Sequence 7841, Application US/09252691B
; GENERAL INFORMATION:
; APPLICANT: Keith G. Weinstock et al.

```

: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROBACTER
: TITLE OF INVENTION: CLOACAE FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196.135
: CURRENT APPLICATION NUMBER: US/09/252,691B
: CURRENT FILING DATE: 1999-02-18
: NUMBER OF SEQ ID NOS: 11324
: SEQ ID NO 7841
: LENGTH: 307
: TYPE: PRT
: ORGANISM: Enterobacter cloacae
: US-09-252-691-7841

```

	Query Match	15.0%	Score 354	DB 16	Length 307
	Best Local Similarity	39.7%	Pred. No. 3,1e-27		
	Matches	85	Conservative	30	Mismatches 69; Indels 10; Gaps 5
QY	226	RAPELKFQVSVKVKAAEDLMDGTDSDLWFGYQOOSHMQIFNGKNSRPFRRVHDYQPEIFLTQ	285		
DB	101	RKDEKFLSLAFLFPWRILGPNISVIGASLYTQKSWMOLNSGESSPFRETNEPOLFLGF	160		
QY	286	PV-YESDLPMDGKVRVIRIGAGVHNSGESAKLSRSNNRPLVLMGEMKUNLTVPRILMGRIE	344		
DB	161	ATDIYEFAGH--TLRQVEVGYNHDSNGRSDPSTRSMNRITTLMAQNGNPMVEYKHH--Y	215		
QY	345	KEGSGSDPDNDLIDYGYGDVRFYLQDENKSNITGVIRYVPRSGKALQIDLYVPLGK	404		
DB	216	VVGS---TDDNPDITRYKMGYOLKRGYOL-GDAVLSAKQOYMMNMTGCGGAELGLSPVPTK	271		
QY	405	GISGFEQIFQGGOSLIDYNNHEATSEFGVGLMND	438		
DB	272	HVRLTYVSGESLIDINFMQIRVGVGLMND	305		

Search completed: November 30, 2001, 14:24:39
Job time: 576 sec

Inis Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:55:42 ; Search time 36.87 Seconds

(Without alignments)
269.771 Million cell updates/sec

Title: US-09-787-083-8

Perfect score: 442
Sequence: 1 MKVSLSTLTLSTLPCFAI...YMHKATSPGVGLMDNMCL 442

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 212252 seqs, 22503292 residues

Word size : 0

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/PTUS.COMB.pep: *
6: /cgn2_6/ptodata/2/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	1.6	16	1	US-07-942-245-98 Sequence 98, Appl
2	7	1.6	552	1	US-08-231-7298-6 Sequence 6, Appl
3	7	1.6	900	4	US-08-890-865A-4 Sequence 4, Appl
4	7	1.6	934	1	US-08-215-805A-80 Sequence 80, Appl
5	7	1.6	1026	2	US-08-342-003-6 Sequence 6, Appl
6	7	1.6	1026	2	US-08-322-760A-6 Sequence 6, Appl
7	6	1.4	10	3	US-08-974-775-30 Sequence 30, Appl
8	6	1.4	11	3	US-08-974-775-29 Sequence 29, Appl
9	6	1.4	12	3	US-08-974-775-28 Sequence 28, Appl
10	6	1.4	13	3	US-08-974-775-8 Sequence 8, Appl
11	6	1.4	13	3	US-08-974-775-27 Sequence 9, Appl
12	6	1.4	14	3	US-08-974-775-10 Sequence 10, Appl
13	6	1.4	14	3	US-08-974-775-32 Sequence 32, Appl
14	6	1.4	15	3	US-08-974-775-5 Sequence 5, Appl
15	6	1.4	16	3	US-08-974-775-4 Sequence 4, Appl
16	6	1.4	20	1	US-08-430-273-1 Sequence 1, Appl
17	6	1.4	20	1	US-07-918-953-13 Sequence 13, Appl
18	6	1.4	21	1	US-07-918-953-15 Sequence 15, Appl
19	6	1.4	21	1	US-08-212-696-1 Sequence 1, Appl
20	6	1.4	21	1	US-08-158-245-1 Sequence 1, Appl
21	6	1.4	21	1	US-08-081-661-13 Sequence 13, Appl
22	6	1.4	21	1	US-08-081-661-15 Sequence 15, Appl
23	6	1.4	21	1	US-08-233-617-1 Sequence 1, Appl
24	6	1.4	21	1	US-08-160-376A-1 Sequence 1, Appl
25	6	1.4	21	1	US-08-304-070-1 Sequence 1, Appl
26	6	1.4	21	1	US-08-285-661-1 Sequence 1, Appl
27	6	1.4	21	1	US-08-285-661-1 Sequence 1, Appl

28	6	1.4	21	1	US-08-301-838-1 Sequence 1, Appl
29	6	1.4	21	1	US-08-389-487-4 Sequence 4, Appl
30	6	1.4	21	1	US-08-389-487-12 Sequence 12, Appl
31	6	1.4	21	1	US-08-507-124-2 Sequence 2, Appl
32	6	1.4	21	1	US-08-507-124-3 Sequence 3, Appl
33	6	1.4	21	1	US-08-342-931-1 Sequence 1, Appl
34	6	1.4	21	1	US-08-400-256-1 Sequence 1, Appl
35	6	1.4	21	2	US-08-508-664-9 Sequence 9, Appl
36	6	1.4	21	2	US-08-353-476-85 Sequence 85, Appl
37	6	1.4	21	2	US-08-353-476-87 Sequence 87, Appl
38	6	1.4	21	2	US-08-484-219-1 Sequence 1, Appl
39	6	1.4	21	2	US-08-979-587-1 Sequence 1, Appl
40	6	1.4	21	2	US-08-992-676-1 Sequence 1, Appl
41	6	1.4	21	2	US-08-992-676-6 Sequence 6, Appl
42	6	1.4	21	2	US-09-134-836-1 Sequence 1, Appl
43	6	1.4	21	2	US-09-134-836-7 Sequence 7, Appl
44	6	1.4	21	3	US-08-967-867-2 Sequence 2, Appl
45	6	1.4	21	4	US-08-932-082-1 Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-07-942-245-98
Sequence 98, Application US/07942245
Patent No. 5639641
GENERAL INFORMATION:
APPLICANT: PEDERSEN, Jan T.
APPLICANT: SEARLE, Stephen M.J.
APPLICANT: REES, Anthony R.
APPLICANT: ROGUSKA, Michael A.
APPLICANT: GUILD, Braydon C.
TITLE OF INVENTION: SURFACE RESIDUE VENERING OF RODENT
TITLE OF INVENTION: ANTIBODIES
NUMBER OF SEQUENCES: 522
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Zimm, Macpeak & Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: HP 9000/700 Workstation
OPERATING SYSTEM: UNIX
SOFTWARE: In house
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/942,245
FILING DATE: 09-SEP-1992
CLASSIFICATION: 530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 98:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-942-245-98

Query Match 1.6%; Score 7; DB 1; Length 16;
Best local Similarity 100.0%; Pred. No. 5.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
QY 387 PRSGGA 393
|||||
Db 4 PRSGGA 10

RESULT 2
US-08-231-729B-6
; Sequence 6, Application US/08231729B
; Patent No. 5618722
; GENERAL INFORMATION:
; APPLICANT: ZENNO, Shuhei
; APPLICANT: SHIRAIISHI, Shinji
; APPLICANT: INOUE, Satoshi
; APPLICANT: SAIGO, Kaoru
; TITLE OF INVENTION: FIRELY LUCIFERASE GENE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LEYDIG, VOIT & MAYER
; STREET: 700 Thirteenth Street, N.W., Suite 300
; City: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,729B
; FILING DATE: 20-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 119050/1993
; FILING DATE: 21-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rose, Herbert C.
; REGISTRATION NUMBER: 29846
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-737-6770
; TELEFAX: 202-737-6776
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 552 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-231-729B-6

Query Match 1.6%; Score 7; DB 1; Length 552;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 LSPFLDR 181
| | | | |
Db 184 LSPFLDR 190

RESULT 3
US-08-890-865A-4
; Sequence 4, Application US/08890865A
; Patent No. 6307019
; GENERAL INFORMATION:
; APPLICANT: Constantini, Franklin
; APPLICANT: Zeng, Li
; TITLE OF INVENTION: AXIN GENE AND USES THEREOF
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; City: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10036
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/890,865A
; FILING DATE: 10-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)278-0400
; TELEFAX: (212)391-0526
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 900 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-890-865A-4

Query Match 1.6%; Score 7; DB 4; Length 900;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 403 GKIGSGY 409
| | | | |
Db 268 GKIGSGY 274

RESULT 4
US-08-215-805A-80
; Sequence 80, Application US/08215805A
; Patent No. 5559008
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; TITLE OF INVENTION: LEUKOTOXIN GENE FROM PASTEURRELLA
; TITLE OF INVENTION: SUIS
; NUMBER OF SEQUENCES: 84
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; City: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/215,805A
; FILING DATE: 22-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Timian, Susan J.
; REGISTRATION NUMBER: 34,103
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1636
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 80:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 934 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein

ORIGINAL SOURCE:
ORGANISM: Pasteurella suis
STRAIN: 5943
IMMEDIATE SOURCE:
LIBRARY: P. suis DNA in Bacteriophage lambda-dash
CLONE: (Lambda)yfc33-37
US-08-215-805A-80

Query Match 1.6%; Score 7; DB 1; Length 934;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 238 VKAEDL 244
|||||||
Db 62 VKAEDL 68

RESULT 5
US-08-542-003-6
Sequence 6, Application US/08542003
Patent No. 5864013
GENERAL INFORMATION:
APPLICANT: Goldberg, Edward B.
TITLE OF INVENTION: MATERIALS FOR THE PRODUCTION OF
TITLE OF INVENTION: NANOMETER STRUCTURES AND USE THEREOF
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennile and Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/542.003
FILING DATE: 13-OCT-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 8471-0005-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: 212-869-8864
TELEX: 66441 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1026 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Bacteriophage T4
IMMEDIATE SOURCE:
CLONE: p37 amino acid
US-08-542-003-6

Query Match 1.6%; Score 7; DB 2; Length 1026;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 128 OTIKGP 134
|||||||
Db 861 OTIKGP 867

RESULT 6
US-08-322-760A-6
Sequence 6, Application US/08322760A
Patent No. 5877279
GENERAL INFORMATION:
APPLICANT: Goldberg, Edward B.
TITLE OF INVENTION: MATERIALS FOR THE PRODUCTION OF
TITLE OF INVENTION: NANOMETER STRUCTURES AND USE THEREOF
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennile and Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/322.760A
FILING DATE: 13-OCT-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 8471-0003-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: 212-869-8864
TELEX: 66441 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1026 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Bacteriophage T4
IMMEDIATE SOURCE:
CLONE: p37 amino acid
US-08-322-760A-6

Query Match 1.6%; Score 7; DB 2; Length 1026;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 128 OTIKGP 134
|||||||
Db 861 OTIKGP 867

RESULT 7
US-08-974-775-30
Sequence 30, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:
APPLICANT: Toback, F. Gary
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
STREET: Plaza Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-974-775-30

```

```

Query Match 1.4%; Score 6; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 424 NHEATS 429
    |||||
DB 5 NHEATS 10

```

```

RESULT 8
US-08-974-775-29
; Sequence 29, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; STREET: Plaza Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid

```

```

; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-974-775-29

```

```

Query Match 1.4%; Score 6; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 424 NHEATS 429
    |||||
DB 5 NHEATS 10

```

```

RESULT 9
US-08-974-775-28
; Sequence 28, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; STREET: Plaza Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-974-775-28

```

```

Query Match 1.4%; Score 6; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 424 NHEATS 429
    |||||
DB 5 NHEATS 10

```

```

RESULT 10
US-08-974-775-8
; Sequence 8, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:

```

```
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; STREET: Plaza Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-974-775-8
```

```
Query Match 1.4%; Score 6; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 424 NHEATS 429
Db 8 NHEATS 13

RESULT 11
US-08-974-775-27
; Sequence 27, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; STREET: Plaza Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
```

```
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-974-775-27
```

```
Query Match 1.4%; Score 6; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 424 NHEATS 429
Db 5 NHEATS 10

RESULT 12
US-08-974-775-9
; Sequence 9, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; STREET: Plaza Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-974-775-9

Query Match 1.4%; Score 6; DB 3; Length 14;
```

Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 424 NHEATS 429
|||||
Db 8 NHEATS 13

RESULT 13
US-08-974-775-10
; Sequence 10, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Clyfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ. ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-974-775-10

Query Match 1.4%; Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 424 NHEATS 429
|||||
Db 8 NHEATS 13

RESULT 14
US-08-974-775-32
; Sequence 32, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE

; STREET: NBC Tower - Suite 3600, 455 N. Clyfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ. ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE: Modified-site
; NAME/KEY: 14
; LOCATION: 14
; OTHER INFORMATION: /product="Ser or Tyr"
US-08-974-775-32

Query Match 1.4%; Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 424 NHEATS 429
|||||
Db 8 NHEATS 13

RESULT 15
US-08-974-775-5
; Sequence 5, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Clyfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:

NAME: Martin, Alice O.
 REGISTRATION NUMBER: 35,601
 REFERENCE/DOCKET NUMBER: 7814/27
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-321-4200
 TELEFAX: 312-321-4299
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 15 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-974-775-5

Query Match 1.4%; Score 6; DB 3; Length 15;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 424 NHEATS 429
 Db 8 NHEATS 13

Search completed: May 9, 2002, 12:55:43
 Job time: 553 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 13:03:24 ; Search time 403.82 Seconds

(Without alignments)
384,464 Million cell updates/sec

Title: US-09-787-083-8

Perfect score: 442
Sequence: 1 MKVSLSTLTLSTLPCFALIA.....YVHETATSGVGLMDNMKL 442

Scoring table:

Gap 60.0 , Gapext 60.0

Searched: 3516493 seqs, 351254056 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3516493

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Pending Patents_AA_Main:*

1: /cgn2_6/prodata/2/paa/PCTUS_COMB.pep.*
2: /cgn2_6/prodata/2/paa/US06_COMB.pep.*
3: /cgn2_6/prodata/2/paa/US07_COMB.pep.*
4: /cgn2_6/prodata/2/paa/US080_COMB.pep.*
5: /cgn2_6/prodata/2/paa/US081_COMB.pep.*
6: /cgn2_6/prodata/2/paa/US082_COMB.pep.*
7: /cgn2_6/prodata/2/paa/US083_COMB.pep.*
8: /cgn2_6/prodata/2/paa/US084_COMB.pep.*
9: /cgn2_6/prodata/2/paa/US085_COMB.pep.*
10: /cgn2_6/prodata/2/paa/US086_COMB.pep.*
11: /cgn2_6/prodata/2/paa/US087_COMB.pep.*
12: /cgn2_6/prodata/2/paa/US088_COMB.pep.*
13: /cgn2_6/prodata/2/paa/US089_COMB.pep.*
14: /cgn2_6/prodata/2/paa/US090_COMB.pep.*
15: /cgn2_6/prodata/2/paa/US091_COMB.pep.*
16: /cgn2_6/prodata/2/paa/US092_COMB.pep.*
17: /cgn2_6/prodata/2/paa/US093_COMB.pep.*
18: /cgn2_6/prodata/2/paa/US094_COMB.pep.*
19: /cgn2_6/prodata/2/paa/US095_COMB.pep.*
20: /cgn2_6/prodata/2/paa/US096_COMB.pep.*
21: /cgn2_6/prodata/2/paa/US097_COMB.pep.*
22: /cgn2_6/prodata/2/paa/US098_COMB.pep.*
23: /cgn2_6/prodata/2/paa/US099_COMB.pep.*
24: /cgn2_6/prodata/2/paa/US60_COMB.pep.*
25: /cgn2_6/prodata/2/paa/US10_MERGED_COMB.pep.*
26: /cgn2_6/prodata/2/paa/US10_MERGED_COMB.pep.*
27: /cgn2_6/prodata/2/paa/US09_MERGED_COMB.pep.*
28: /cgn2_6/prodata/2/paa/US08_MERGED_COMB.pep.*
29: /cgn2_6/prodata/2/paa/US08_MERGED_COMB.pep.*
30: /cgn2_6/prodata/2/paa/US06_MERGED_COMB.pep.*
31: /cgn2_6/prodata/2/paa/US06_MERGED_COMB.pep.*
32: /cgn2_6/prodata/2/paa/PCT_MERGED_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	442	100.0	442	21	US-09-787-083-8

2	327	74.0	442	21	US-09-787-083-2	Sequence 2, Appl
3	316	71.5	442	21 <th>US-09-787-083-6</th> <th>Sequence 6, Appl</th>	US-09-787-083-6	Sequence 6, Appl
4	289	65.4	442	21 <th>US-09-787-083-4</th> <th>Sequence 3202, Ap</th>	US-09-787-083-4	Sequence 3202, Ap
5	215	48.6	346	19 <th>US-09-540-236-3202</th> <th>Sequence 3019, Ap</th>	US-09-540-236-3202	Sequence 3019, Ap
6	211	47.7	375	24 <th>US-60-128-476-3019</th> <th>Sequence 3197, Ap</th>	US-60-128-476-3019	Sequence 3197, Ap
7	102	23.1	109	9 <th>US-09-540-236-3197</th> <th>Sequence 3016, Ap</th>	US-09-540-236-3197	Sequence 3016, Ap
8	102	23.1	132	24 <th>US-60-128-476-3016</th> <th>Sequence 48393, A</th>	US-60-128-476-3016	Sequence 48393, A
9	9	2.0	255	1 <th>PCT-US01-08631-48393</th> <th>Sequence 4, Appl</th>	PCT-US01-08631-48393	Sequence 4, Appl
10	9	2.0	374	21 <th>US-09-787-084-4</th> <th>Sequence 2, Appl</th>	US-09-787-084-4	Sequence 2, Appl
11	9	2.0	375	21 <th>US-09-787-084-2</th> <th>Sequence 1681, Ap</th>	US-09-787-084-2	Sequence 1681, Ap
12	8	1.8	227	25 <th>US-60-361-742-1681</th> <th>Sequence 8021, Ap</th>	US-60-361-742-1681	Sequence 8021, Ap
13	8	1.8	297	18 <th>US-09-489-039A-8021</th> <th>Sequence 8, Appl</th>	US-09-489-039A-8021	Sequence 8, Appl
14	7	1.6	50	15 <th>US-09-161-939-8</th> <th>Sequence 674, App</th>	US-09-161-939-8	Sequence 674, App
15	7	1.6	50	15 <th>US-09-161-939A-8</th> <th>Sequence 1266, Ap</th>	US-09-161-939A-8	Sequence 1266, Ap
16	7	1.6	52	21 <th>US-09-760-498-674</th> <th>Sequence 1001, Ap</th>	US-09-760-498-674	Sequence 1001, Ap
17	7	1.6	58	24 <th>US-60-170-429-1266</th> <th>Sequence 1580, Ap</th>	US-60-170-429-1266	Sequence 1580, Ap
18	7	1.6	58	24 <th>US-60-182-568-1001</th> <th>Sequence 23275, A</th>	US-60-182-568-1001	Sequence 23275, A
19	7	1.6	65	19 <th>US-09-595-329A-1580</th> <th>Sequence 4265, Ap</th>	US-09-595-329A-1580	Sequence 4265, Ap
20	7	1.6	70	18 <th>US-09-417-507-23275</th> <th>Sequence 4265, Ap</th>	US-09-417-507-23275	Sequence 4265, Ap
21	7	1.6	73	24 <th>US-60-160-203-4820</th> <th>Sequence 6621, Ap</th>	US-60-160-203-4820	Sequence 6621, Ap
22	7	1.6	73	24 <th>US-60-162-247-4265</th> <th>Sequence 1013, Ap</th>	US-60-162-247-4265	Sequence 1013, Ap
23	7	1.6	73	24 <th>US-60-169-840-6621</th> <th>Sequence 1013, Ap</th>	US-60-169-840-6621	Sequence 1013, Ap
24	7	1.6	75	1 <th>PCT-US01-01340-1013</th> <th>Sequence 1013, Ap</th>	PCT-US01-01340-1013	Sequence 1013, Ap
25	7	1.6	75	21 <th>US-09-764-869-1013</th> <th>Sequence 41947, A</th>	US-09-764-869-1013	Sequence 41947, A
26	7	1.6	75	26 <th>US-10-091-504-1013</th> <th>Sequence 4426, Ap</th>	US-10-091-504-1013	Sequence 4426, Ap
27	7	1.6	85	24 <th>US-60-188-162-4426</th> <th>Sequence 41947, A</th>	US-60-188-162-4426	Sequence 41947, A
28	7	1.6	91	27 <th>US-09-708-427-41947</th> <th>Sequence 888, App</th>	US-09-708-427-41947	Sequence 888, App
29	7	1.6	99	24 <th>US-60-187-387-1177</th> <th>Sequence 72, Appl</th>	US-60-187-387-1177	Sequence 72, Appl
30	7	1.6	107	27 <th>US-09-594-595B-888</th> <th>Sequence 100, Appl</th>	US-09-594-595B-888	Sequence 100, Appl
31	7	1.6	116	1 <th>PCT-US01-01384-72</th> <th>Sequence 1953, Ap</th>	PCT-US01-01384-72	Sequence 1953, Ap
32	7	1.6	116	1 <th>PCT-US01-01384-99</th> <th>Sequence 1954, Ap</th>	PCT-US01-01384-99	Sequence 1954, Ap
33	7	1.6	116	1 <th>PCT-US01-01384-100</th> <th>Sequence 1954, Ap</th>	PCT-US01-01384-100	Sequence 1954, Ap
34	7	1.6	116	1 <th>PCT-US01-11988-1953</th> <th>Sequence 1954, Ap</th>	PCT-US01-11988-1953	Sequence 1954, Ap
35	7	1.6	116	1 <th>PCT-US01-11988-1954</th> <th>Sequence 1953, Ap</th>	PCT-US01-11988-1954	Sequence 1953, Ap
36	7	1.6	116	1 <th>PCT-US01-11988-1955</th> <th>Sequence 1953, Ap</th>	PCT-US01-11988-1955	Sequence 1953, Ap
37	7	1.6	116	22 <th>US-09-833-245-1953</th> <th>Sequence 1954, Ap</th>	US-09-833-245-1953	Sequence 1954, Ap
38	7	1.6	116	22 <th>US-09-833-245-1954</th> <th>Sequence 1954, Ap</th>	US-09-833-245-1954	Sequence 1954, Ap
39	7	1.6	116	22 <th>US-09-833-245-1955</th> <th>Sequence 1954, Ap</th>	US-09-833-245-1955	Sequence 1954, Ap
40	7	1.6	133	1 <th>PCT-US01-08631-57915</th> <th>Sequence 13005, A</th>	PCT-US01-08631-57915	Sequence 13005, A
41	7	1.6	134	1 <th>PCT-US01-14827-13005</th> <th>Sequence 12780, A</th>	PCT-US01-14827-13005	Sequence 12780, A
42	7	1.6	158	18 <th>US-09-489-039A-12780</th> <th>Sequence 42, Appl</th>	US-09-489-039A-12780	Sequence 42, Appl
43	7	1.6	158	21 <th>US-09-138-132-42</th> <th>Sequence 12847, A</th>	US-09-138-132-42	Sequence 12847, A
44	7	1.6	158	21 <th>US-09-789-919-42</th> <th></th>	US-09-789-919-42	
45	7	1.6	160	18 <th>US-09-489-039A-12847</th> <th></th>	US-09-489-039A-12847	

ALIGNMENTS

RESULT 1
US-09-787-083-8
Sequence 8, Application US/09787083
GENERAL INFORMATION:
APPLICANT: Ruellie, Jean-Louis
TITLE OF INVENTION: Novel Compounds
FILE REFERENCE: BM45332
CURRENT APPLICATION NUMBER: US/09/787, 083
CURRENT FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: PCT/EP99/06781
PRIOR FILING DATE: 1999-09-14
PRIOR APPLICATION NUMBER: GB 9820002.5
PRIOR FILING DATE: 1998-09-14
NUMBER OF SEQ ID NOS: 13
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 442
TYPE: PRT
ORGANISM: Moraxella catarrhalis
US-09-787-083-8

Query Match 100.0%; Score 442; DB 21; Length 442;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVSLSTLTLILPCFAILAIQQAQAVNPVAFVDEVRENDLGODNELPIDVOSATQSA 60
DB 1 MKVSLSTLTLILPCFAILAIQQAQAVNPVAFVDEVRENDLGODNELPIDVOSATQSA 60
QY 61 STDTANPLDEHEBELYTTALENTMLINCALNODIMRLACYDTLVHGETPAVYIKTKRSI 120
DB 61 STDTANPLDEHEBELYTTALENTMLINCALNODIMRLACYDTLVHGETPAVYIKTKRSI 120
QY 121 RLDETTMOTIKGRPOVYVOETTDPIFLMGNEKMLTKKDAKOLEYAKOFTPLSLDFLD 180
DB 121 RLDETTMOTIKGRPOVYVOETTDPIFLMGNEKMLTKKDAKOLEYAKOFTPLSLDFLD 180
QY 181 RNNTPLMSSRPHNPMYVLPFIEMHGKPNRSPNTPSHEAKOFTPNEFRAPELKFVSVKVA 240
DB 181 RNNTPLMSSRPHNPMYVLPFIEMHGKPNRSPNTPSHEAKOFTPNEFRAPELKFVSVKVA 240
QY 241 AEDLMGTDSLWMFGYTOQSHMOIFNGKNSRPFVHDYQPEIFLTOPYSDLPMDGKVRMI 300
DB 241 AEDLMGTDSLWMFGYTOQSHMOIFNGKNSRPFVHDYQPEIFLTOPYSDLPMDGKVRMI 300
QY 301 GMGAVHHSNCSAKLSRSNRRAYILAMGEMKNLTVMPRIGRIFFKESGSGQPDNDPILD 360
DB 301 GMGAVHHSNCSAKLSRSNRRAYILAMGEMKNLTVMPRIGRIFFKESGSGQPDNDPILD 360
QY 361 YYGVDVRFPLYOLENKSNIIGTVRNPBSGKALQLDYVPLGKISGYFOIFOGYGQSL 420
DB 361 YYGVDVRFPLYOLENKSNIIGTVRNPBSGKALQLDYVPLGKISGYFOIFOGYGQSL 420
QY 421 IDYNHEATSGVGLMNDMMGL 442
DB 421 IDYNHEATSGVGLMNDMMGL 442

RESULT 2
US-09-787-083-2
; Sequence 2, Application US/09787083
; GENERAL INFORMATION:
; APPLICANT: Ruelle, Jean-Louis
; TITLE OF INVENTION: Novel Compounds
; FILE REFERENCE: BM45332
; CURRENT APPLICATION NUMBER: US/09/787,083
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: PCT/EP99/06781
; PRIOR FILING DATE: 1999-09-14
; PRIOR APPLICATION NUMBER: GB 9820002.5
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 442
; TYPE: PRF
; ORGANISM: Moraxella catarrhalis
US-09-787-083-2

Query Match 74.0%; Score 327; DB 21; Length 442;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 15 CFAILAIQQAQAVNPVAFVDEVRENDLGODNELPIDVOSATQSASTDTANPLDEHEPE 74
DB 15 CFAILAIQQAQAVNPVAFVDEVRENDLGODNELPIDVOSATQSASTDTANPLDEHEPE 74
QY 75 LYTALLENKMLINCALNODIMRLACYDTLVHGETPAVYIKTKRSIRLDETTMOTIKGRP 134
DB 75 LYTALLENKMLINCALNODIMRLACYDTLVHGETPAVYIKTKRSIRLDETTMOTIKGRP 134
QY 135 QVVYQETTDPIFLMGNEKMLTKKDAKOLEYAKOFTPLSLDFLDNRNNPLWSSRPHN 194
DB 135 QVVYQETTDPIFLMGNEKMLTKKDAKOLEYAKOFTPLSLDFLDNRNNPLWSSRPHN 194

QY 195 MYVLPFIEMHGKPNRSPNTPSHEAKOFTPNEFRAPELKFVSVKVAEDLMGTDSLWMFG 254
DB 195 MYVLPFIEMHGKPNRSPNTPSHEAKOFTPNEFRAPELKFVSVKVAEDLMGTDSLWMFG 254
QY 255 YTOQSHMOIFNGKNSRPFVHDYQPEIFLTOPYSDLPMDGKVRMIGMGAVHHSNCSAK 314
DB 255 YTOQSHMOIFNGKNSRPFVHDYQPEIFLTOPYSDLPMDGKVRMIGMGAVHHSNCSAK 314
QY 315 LSRSMNRRAYILAMGEMKNLTVMPRIGRIFFKESGSGQPDNDPILDYGYGVDVRFPLYOLE 374
DB 315 LSRSMNRRAYILAMGEMKNLTVMPRIGRIFFKESGSGQPDNDPILDYGYGVDVRFPLYOLE 374
QY 375 NKSNIIGTVRNPBSGKALQLDYVPLGKISGYFOIFOGYGQSLIDYNHEATSGVGL 434
DB 375 NKSNIIGTVRNPBSGKALQLDYVPLGKISGYFOIFOGYGQSLIDYNHEATSGVGL 434
QY 435 MLDMMGL 442
DB 435 MLDMMGL 442

RESULT 3
US-09-787-083-6
; Sequence 6, Application US/09787083
; GENERAL INFORMATION:
; APPLICANT: Ruelle, Jean-Louis
; TITLE OF INVENTION: Novel Compounds
; FILE REFERENCE: BM45332
; CURRENT APPLICATION NUMBER: US/09/787,083
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: PCT/EP99/06781
; PRIOR FILING DATE: 1999-09-14
; PRIOR APPLICATION NUMBER: GB 9820002.5
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 442
; TYPE: PRF
; ORGANISM: Moraxella catarrhalis
US-09-787-083-6

Query Match 71.5%; Score 316; DB 21; Length 442;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 26 AVNPVAFVDEVRENDLGODNELPIDVOSATQSASTDTANPLDEHEBELYTTALENTKM 85
DB 26 AVNPVAFVDEVRENDLGODNELPIDVOSATQSASTDTANPLDEHEBELYTTALENTKM 85
QY 86 LINCALNODIMRLACYDTLVHGETPAVYIKTKRSIRLDETTMOTIKGRPOVYQETTDPI 145
DB 86 LINCALNODIMRLACYDTLVHGETPAVYIKTKRSIRLDETTMOTIKGRPOVYQETTDPI 145
QY 146 FLMGNEKMLTKKDAKOLEYAKOFTPLSLDFLDNRNNPLWSSRPHNPMYVLPFIEMHGK 205
DB 146 FLMGNEKMLTKKDAKOLEYAKOFTPLSLDFLDNRNNPLWSSRPHNPMYVLPFIEMHGK 205
QY 206 PNRSPNTPSHEAKOFTPNEFRAPELKFVSVKVAEDLMGTDSLWMFGYTOQSHMOIN 265
DB 206 PNRSPNTPSHEAKOFTPNEFRAPELKFVSVKVAEDLMGTDSLWMFGYTOQSHMOIN 265
QY 266 GKNRSPRFRVHDYQPEIFLTOPYSDLPMDGKVRMIGMGAVHHSNCSAKLSRSNRRAYIL 325
DB 266 GKNRSPRFRVHDYQPEIFLTOPYSDLPMDGKVRMIGMGAVHHSNCSAKLSRSNRRAYIL 325
QY 326 AGEMKNLTVMPRIGRIFFKESGSGQPDNDPILDYGYGVDVRFPLYOLENKSNIIGTVR 385
DB 326 AGEMKNLTVMPRIGRIFFKESGSGQPDNDPILDYGYGVDVRFPLYOLENKSNIIGTVR 385
QY 386 NPSRSGKALQLDYVPLGKISGYFOIFOGYGQSLIDYNHEATSGVGLMNDMMGL 442
DB 386 NPSRSGKALQLDYVPLGKISGYFOIFOGYGQSLIDYNHEATSGVGLMNDMMGL 442

Db 386 NPSRGKALQIDYVYPLGKIGSYFQIFQYGGSLIDYNHEATSGVGLMLNDMKG 442

RESULT 4
US-09-787-083-4
: Sequence 4, Application US/09787083
: GENERAL INFORMATION:
: APPLICANT: Ruelle, Jean-Louis
: TITLE OF INVENTION: Novel Compounds
: FILE REFERENCE: BM45332
: CURRENT APPLICATION NUMBER: US/09/787, 083
: PRIOR FILING DATE: 2001-03-14
: PRIOR APPLICATION NUMBER: PCT/EP99/06781
: PRIOR FILING DATE: 1999-09-14
: PRIOR APPLICATION NUMBER: GB 9820002.5
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 4
: LENGTH: 442
: TYPE: PRT
: ORGANISM: Moraxella catarrhalis
US-09-787-083-4

Query Match 65.4%; Score 289; DB 21; Length 442;
Best Local Similarity 99.7%; Pred. No. 2.3e-291;
Matches 389; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 53 VQATQASATDTPANPLDEHEPELYTTALENTMLINCALNODIMRLACYPDVLHGETPA 112
DB 53 VQATQASATDTPANPLDEHEPELYTTALENTMLINCALNODIMRLACYPDVLHGETPA 112
QY 113 VIKTKRSIRLDETTMOTIKGRQVYVYQETDPIFLMGNEKMLTKKAKOLEYAKOPTP 172
DB 113 VIKTKRSIRLDETTMOTIKGRQVYVYQETDPIFLMGNEKMLTKKAKOLEYAKOPTP 172
QY 173 LSLFDDLRNNTPLMSSRPNRPVYLPFIEMGKRPKRSPTNSHEKAKOPTPEFRAPELKE 232
DB 173 LSLFDDLRNNTPLMSSRPNRPVYLPFIEMGKRPKRSPTNSHEKAKOPTPEFRAPELKE 232
QY 233 QVSKVKAEDLMGTDSDLMFGYTOQSHMOJFNKNSRPFRVNDYQPIFLTOPIYSDLR 292
DB 233 QVSKVKAEDLMGTDSDLMFGYTOQSHMOJFNKNSRPFRVNDYQPIFLTOPIYSDLR 292
QY 293 WDGKVRMIGMGAHVHNSGESAKLSRSNNRAYLTMGMEMKNTLVMPRIWGRIEKGSGSQP 352
DB 293 WDGKVRMIGMGAHVHNSGESAKLSRSNNRAYLTMGMEMKNTLVMPRIWGRIEKGSGSQP 352
QY 353 DDNDPDIIDYGYGVRFYLOLEKNSISGTVRPNRSGKALQIDYVYPLGKIGSYFOI 412
DB 353 DDNDPDIIDYGYGVRFYLOLEKNSISGTVRPNRSGKALQIDYVYPLGKIGSYFOI 412
QY 413 FQYGGQSLIDYNHEATSGVGLMLNDMKG 442
DB 413 FQYGGQSLIDYNHEATSGVGLMLNDMKG 442

RESULT 5
US-09-540-236-3202
: Sequence 3202, Application US/09540236
: GENERAL INFORMATION:
: APPLICANT: Gary L. Breton et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
: FILE REFERENCE: 2709, 2005-001
: CURRENT APPLICATION NUMBER: US/09/540, 236
: NUMBER OF SEQ ID NOS: 3840
: SEQ ID NO 3202
: LENGTH: 346
: TYPE: PRT
: ORGANISM: M. catarrhalis

US-09-540-236-3202

Query Match 48.6%; Score 215; DB 19; Length 346;
Best Local Similarity 99.7%; Pred. No. 2.1e-214;
Matches 315; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 26 AVPNPVAFVDEVSENDIGDQNELEPIDVQATQASATDTPANPLDEHEPELYTTALENTM 85
DB 26 AVPNPVAFVDEVSENDIGDQNELEPIDVQATQASATDTPANPLDEHEPELYTTALENTM 85
QY 86 LINCASALNODIMRLACYPDVLHGETPAVYIKTKRSIRLDETTMOTIKGRQVYVYQETDPI 145
DB 86 LINCASALNODIMRLACYPDVLHGETPAVYIKTKRSIRLDETTMOTIKGRQVYVYQETDPI 145
QY 146 FLMGNEKMLTKKAKOLEYAKOPTPLSLFDDLRNNTPLMSSRPNRPVYLPFIEMGK 205
DB 146 FLMGNEKMLTKKAKOLEYAKOPTPLSLFDDLRNNTPLMSSRPNRPVYLPFIEMGK 205
QY 206 PNRSPNTPSHAKOPTPEFRAPELKFOYSVKVAAEDLMGTDSDLMFGYTOQSHMOJFN 265
DB 206 PNRSPNTPSHAKOPTPEFRAPELKFOYSVKVAAEDLMGTDSDLMFGYTOQSHMOJFN 265
QY 268 GKNRSPFRVNDYQPIFLTOPIYSDLRPMWDGKVRMIGMGAHVHNSGESAKLSRSNNRAYL 327
DB 268 GKNRSPFRVNDYQPIFLTOPIYSDLRPMWDGKVRMIGMGAHVHNSGESAKLSRSNNRAYL 327
QY 326 AGMEMKNTLVMPRIWG 341
DB 326 AGMEMKNTLVMPRIWG 341

RESULT 6
US-60-128-476-3019
: Sequence 3019, Application US/60128476
: GENERAL INFORMATION:
: APPLICANT: Gary L. Breton et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196, 139
: CURRENT APPLICATION NUMBER: US/60/128, 476
: NUMBER OF SEQ ID NOS: 5002
: SEQ ID NO 3019
: LENGTH: 375
: TYPE: PRT
: ORGANISM: Moraxella catarrhalis
US-60-128-476-3019

Query Match 47.7%; Score 211; DB 24; Length 375;
Best Local Similarity 99.7%; Pred. No. 3.3e-210;
Matches 311; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 26 AVPNPVAFVDEVSENDIGDQNELEPIDVQATQASATDTPANPLDEHEPELYTTALENTM 85
DB 26 AVPNPVAFVDEVSENDIGDQNELEPIDVQATQASATDTPANPLDEHEPELYTTALENTM 85
QY 86 LINCASALNODIMRLACYPDVLHGETPAVYIKTKRSIRLDETTMOTIKGRQVYVYQETDPI 145
DB 86 LINCASALNODIMRLACYPDVLHGETPAVYIKTKRSIRLDETTMOTIKGRQVYVYQETDPI 145
QY 146 FLMGNEKMLTKKAKOLEYAKOPTPLSLFDDLRNNTPLMSSRPNRPVYLPFIEMGK 205
DB 146 FLMGNEKMLTKKAKOLEYAKOPTPLSLFDDLRNNTPLMSSRPNRPVYLPFIEMGK 205
QY 206 PNRSPNTPSHAKOPTPEFRAPELKFOYSVKVAAEDLMGTDSDLMFGYTOQSHMOJFN 265
DB 206 PNRSPNTPSHAKOPTPEFRAPELKFOYSVKVAAEDLMGTDSDLMFGYTOQSHMOJFN 265
QY 268 GKNRSPFRVNDYQPIFLTOPIYSDLRPMWDGKVRMIGMGAHVHNSGESAKLSRSNNRAYL 327
DB 268 GKNRSPFRVNDYQPIFLTOPIYSDLRPMWDGKVRMIGMGAHVHNSGESAKLSRSNNRAYL 327

QY 326 AGMEMKNTVMP 337
Db 328 AGMEMKNTVMP 339

RESULT 7
US-09-540-236-3197

; Sequence 3197, Application US/09540236
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALLIS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3197
; LENGTH: 109
; TYPE: PRT
; ORGANISM: M.catarrhalls
US-09-540-236-3197

Query Match 23.1%; Score 102; DB 19; Length 109;
Best Local Similarity 100.0%; Pred. No. 3.3e-97;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 GRIFKESGSDPDNPILDYGGVRFYLOENKSNISGTVARNRSGKALQLDYVY 400
Db 8 GRIFKESGSDPDNPILDYGGVRFYLOENKSNISGTVARNRSGKALQLDYVY 67
QY 401 PLKGISGYFOIFQGYQSLIDYNHEATSFVGMLNDMGL 442
Db 68 PLKGISGYFOIFQGYQSLIDYNHEATSFVGMLNDMGL 109

RESULT 8
US-60-128-476-3016

; Sequence 3016, Application US/60128476
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALLIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.139
; CURRENT APPLICATION NUMBER: US/60/128,476
; CURRENT FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 5002
; SEQ ID NO 3016
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Moraxella catarrhalls
US-60-128-476-3016

Query Match 23.1%; Score 102; DB 24; Length 132;
Best Local Similarity 100.0%; Pred. No. 3.9e-97;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 GRIFKESGSDPDNPILDYGGVRFYLOENKSNISGTVARNRSGKALQLDYVY 400
Db 31 GRIFKESGSDPDNPILDYGGVRFYLOENKSNISGTVARNRSGKALQLDYVY 90
QY 401 PLKGISGYFOIFQGYQSLIDYNHEATSFVGMLNDMGL 442
Db 91 PLKGISGYFOIFQGYQSLIDYNHEATSFVGMLNDMGL 132

RESULT 9
PCT-US01-08631-48393
; Sequence 48393, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc

; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 48393

; LENGTH: 255
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (135)..(152)
; OTHER INFORMATION: ArgE / dape / ACY1 / CPE2 / yacs family proteins domain
; OTHER INFORMATION: Identified by eMATRIX, accession number BL00738B, p-value=1.96
; NAME/KEY: DOMAIN
; LOCATION: (2)..(125)
; OTHER INFORMATION: Peptidase family M20/M25/M40 domain identified by Pfam,
; NAME/KEY: misc.feature
; LOCATION: (1)..(255)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
PCT-US01-08631-48393

Query Match 2.0%; Score 9; DB 1; Length 255;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIFLTQPV 287
Db 204 PEIFLTQPV 212

RESULT 10
US-09-787-084-4

; Sequence 4, Application US/09787084
; GENERAL INFORMATION:
; APPLICANT: Ruelie, Jean-Louis
; TITLE OF INVENTION: Polynucleotides and Polypeptides BASB033
; TITLE OF INVENTION: from Neisseria Meningitidis and Their Uses
; FILE REFERENCE: BM45331
; CURRENT APPLICATION NUMBER: US/09/787,084
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: PCT/EP99/06718
; PRIOR FILING DATE: 1999-09-09
; PRIOR APPLICATION NUMBER: GB 9820003.3
; PRIOR FILING DATE: 1998-09-14
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-787-084-4

Query Match 2.0%; Score 9; DB 21; Length 374;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIFLTQPV 287
Db 213 PEIFLTQPV 221

RESULT 11
US-09-787-084-2

```
; Sequence 2, Application US/09787084
; GENERAL INFORMATION:
; APPLICANT: Rueille, Jean-Louis
; TITLE OF INVENTION: Polynucleotides and Polypeptides BASB033
; TITLE OF INVENTION: from Neisseria Meningitidis and Their Uses
; FILE REFERENCE: BM45331
; CURRENT APPLICATION NUMBER: US/09/787,084
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: PCT/EP99/06718
; PRIOR FILING DATE: 1999-09-09
; PRIOR APPLICATION NUMBER: GB 9820003.3
; PRIOR FILING DATE: 1998-09-14
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-787-084-2
```

```
Query Match          2.0%; Score 9; DB 21; Length 375;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 279 PEIFLTQPV 287
    |||||||
DB 214 PEIFLTQPV 222
```

```
RESULT 12
US-60-361-742-1681
; Sequence 1681, Application US/60361742
; GENERAL INFORMATION:
; APPLICANT: Fidelity Systems, Inc.
; APPLICANT: Slesarev, Alexei I.
; APPLICANT: Mezhevaya, Katja V.
; APPLICANT: Polushin, Nikolai N.
; APPLICANT: Shcherbina, Olga V.
; APPLICANT: Shakhova, Vera V.
; APPLICANT: Malykh, Andrei G.
; APPLICANT: Kozavkin, Sergei A.
; TITLE OF INVENTION: Fidelity Systems, Inc.
; FILE REFERENCE: New
; CURRENT APPLICATION NUMBER: US/60/361,742
; CURRENT FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 1692
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1681
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Methanopyrus kandleri
US-60-361-742-1681
```

```
Query Match          1.8%; Score 8; DB 25; Length 227;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 105 LVHGETPA 112
    |||||||
DB 106 LVHGETPA 113
```

```
RESULT 13
US-09-489-039A-8021
; Sequence 8021, Application US/09489039A
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
```

```
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8021
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8021
```

```
Query Match          1.8%; Score 8; DB 18; Length 297;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 431 GVGIMLND 438
    |||||||
DB 288 GVGIMLND 295
```

```
RESULT 14
US-09-161-939-8
; Sequence 8, Application US/09161939
; GENERAL INFORMATION:
; APPLICANT: Shinkets, Richard A.
; TITLE OF INVENTION: Genes and Proteins Predictive and Therapeutic for
; TITLE OF INVENTION: Stroke, Hypertension, Diabetes, and Obesity
; FILE REFERENCE: 15966-527 SHR genes and proteins
; CURRENT APPLICATION NUMBER: US/09/161,939
; CURRENT FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-161-939-8
```

```
Query Match          1.6%; Score 7; DB 15; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 114 IKTKRSI 120
    |||||||
DB 22 IKTKRSI 28
```

```
RESULT 15
US-09-161-939A-8
; Sequence 8, Application US/09161939A
; GENERAL INFORMATION:
; APPLICANT: Shinkets, Richard A.
; TITLE OF INVENTION: Genes and Proteins Predictive and Therapeutic for
; TITLE OF INVENTION: Stroke, Hypertension, Diabetes, and Obesity
; FILE REFERENCE: 15966-527
; CURRENT APPLICATION NUMBER: US/09/161,939A
; CURRENT FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-161-939A-8
```

```
Query Match          1.6%; Score 7; DB 15; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 114 IKTKRSI 120
    |||||||
```

Thu May 9 13:50:35 2002

us-09-787-083-8.qlt.rapm

Page 6

Db 22 IKTksi 28

Search completed: May 9, 2002, 13:03:25
Job time: 615 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:56:33 ; Search time 44.37 Seconds
(without alignments)
758.627 Million cell updates/sec

Title: US-09-787-083-6

Perfect score: 442
Sequence: 1 MKVSLSTLTLSTLSCFAILA.....YNHEATSPGVGLMDNMGL 442

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: PIR-68:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	2.0	382	2	EB1195 phospholipase A1,
2	9	2.0	409	2	EB1831 probable phospholip
3	8	1.8	286	2	B36971 outer membrane pho
4	8	1.8	299	2	T47989 RAV-like protein -
5	8	1.8	1396	2	S36851 L-shaped tail fibre
6	7	1.6	100	2	S61050 probable membrane
7	7	1.6	116	2	F86676 hypothetical prote
8	7	1.6	122	2	E75377 hypothetical prote
9	7	1.6	158	2	E86498 pts IIA protein [1
10	7	1.6	197	2	E72124 tubulin alpha-1 ch
11	7	1.6	219	2	T19438 hypothetical prote
12	7	1.6	237	2	C84976 tRNA (guanine-N1-)
13	7	1.6	240	2	A61544 tubulin alpha chai
14	7	1.6	264	2	A44959 coat protein - pot
15	7	1.6	267	2	A60366 coat protein - pot
16	7	1.6	267	2	S14001 genome polypotein
17	7	1.6	267	2	S26630 capsid protein - p
18	7	1.6	267	2	JC1527 coat protein - pot
19	7	1.6	267	2	S13239 tail fiber protein
20	7	1.6	270	2	S73734 abc transport ATP
21	7	1.6	273	2	C72338 transaminase B hom
22	7	1.6	287	2	A72421 oligopeptide ABC t
23	7	1.6	288	2	T36237 probable phosphotr
24	7	1.6	306	2	T27985 hypothetical prote
25	7	1.6	310	2	T02200 probable RAV-like
26	7	1.6	313	2	JT0960 polypotein - pota
27	7	1.6	321	2	H82240 conserved hypothet
28	7	1.6	324	2	T23876 hypothetical prote

30	7	1.6	326	2	A43939 proteinase inhibitor
31	7	1.6	338	2	F69437 hypothetical prote
32	7	1.6	341	2	T47653 pectate lyase-like
33	7	1.6	342	2	G64411 hypothetical prote
34	7	1.6	342	2	C64394 hypothetical prote
35	7	1.6	347	1	DNHUN2 NADH dehydrogenase
36	7	1.6	347	2	T11051 NADH dehydrogenase
37	7	1.6	347	2	B59153 NADH dehydrogenase
38	7	1.6	349	2	T43920 yfuc protein [impo
39	7	1.6	370	2	G83219 probable dihydrol
40	7	1.6	379	2	S13556 genome polypotein
41	7	1.6	381	2	G82132 ribonuclease D VC1
42	7	1.6	382	2	S13237 tail fiber protein
43	7	1.6	384	2	S70638 tubulin alpha chai
44	7	1.6	384	2	S70641 tubulin alpha chai
45	7	1.6	384	2	S70640 tubulin alpha chai

ALIGNMENTS

RESULT 1

EB1195 phospholipase A1, probable NMB0464 [imported] - Neisseria meningitidis (strain MC58 s

C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C:Accession: EB1195

R:Retelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.

ri, H.; Qin, H.; Yamathayan, J.; Gall, J.; Scarlato, V.; Maignani, V.; Piazza, M.

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;

A>Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: AB1000; MUID:2017555

A:Accession: EB1195

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-382 <TEP>

A:Cross-references: GB:AE002403; GB:AE002098; NID:g7225688; PIDN:AAFA0901.1; PID:g7222

A:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: NMB0464

Query Match 2.0%; Score 9; DB 2; Length 382;

Best Local Similarity 100.0%; Pred. No. 0.46;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIFLTPY 287

DB 221 PEIFLTPY 229

RESULT 2

H81831 phospholipase NMA2021 [imported] - Neisseria meningitidis (strain Z2491 serog

C:Species: Neisseria meningitidis

C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C:Accession: H81831

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo

lloyd, S.; Jørgensen, K.; Leather, S.; Mouton, S.; Mungall, K.; Quail, M.A.; Rajandre

Nature 404, 502-506, 2000

A>Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491

A:Reference number: AB1775; MUID:2022556

A:Accession: H81831

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-409 <PAR>

A:Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CA885240.1; PID:g738

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: NMA2021

Query Match 2.0%; Score 9; DB 2; Length 409;
Best Local Similarity 100.0%; Pred. No. 0.49;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIEFLTPV 287
|||||
DB 248 PEIEFLTPV 256

RESULT 3
B36971

Outer membrane phospholipase A (EC 3.1.1.-) precursor - *Klebsiella pneumoniae*
C:Species: *Klebsiella pneumoniae*
C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 18-Jun-1999
C:Accession: B36971; S40129
R:Brok, R.G.P.M.; Brinkman, E.; van Boxtel, R.; Bekkers, A.C.A.P.A.; Verheij, H.M.; Tomm
J. Bacteriol 176, 861-870, 1994
A:Title: Molecular characterization of enterobacterial pldA genes encoding outer membran
A:Reference number: A36971; MUID:94131966
A:Accession: B36971
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-286 <PRO>
A:Cross-references: EMBL:X76901; NID:g436880; PIDN:CAA54223.1; PID:g436881
A:Note: authors translated the codon AAG for residue 112 as Arg
C:Genetics:
A:Gene: pldA
C:Superfamily: bacterial phospholipase A1
C:Keywords: carboxylic ester hydrolase

Query Match 1.8%; Score 8; DB 2; Length 286;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 431 GVGMLND 438
|||||
DB 277 GVGMLND 284

RESULT 4
T47989

RAV-like protein - *Arabidopsis thaliana*
N:Alternate names: protein F21F14.140
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47989
R:Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Catolico, L.; Artiguenave, F.; Sa
submitted to the Protein Sequence Database, February 2000
A:Reference number: Z24481
A:Accession: T47989
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-299 <CHO>
A:Cross-references: EMBL:AL138642
A:Experimental source: cultivar Columbia; BAC clone F21F14
C:Genetics:
A:Map position: 3
A:Note: F21F14.140

Query Match 1.8%; Score 8; DB 2; Length 299;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 SLSFDDR 181
|||||
DB 292 SLSFDDR 299

RESULT 5
S36851

L-shaped tail fiber protein - phage T5
N:Alternate names: ltf protein
C:Species: phage T5
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 17-Nov-2000
C:Accession: S65934; S01984; S36851
R:Kallman, A.V.; Kulshin, V.E.; Shlyupnikov, M.G.; Ksenzenko, V.N.; Kiyukov, V.M.
FEBS Lett. 366, 46-48, 1995
A:Title: The nucleotide sequence of the bacteriophage T5 ltf gene.
A:Reference number: S65934; MUID:95309401
A:Accession: S65934
A:Molecule type: DNA

A:Residues: 11396 <KAL>
A:Cross-references: EMBL:X69460; NID:g15415; PIDN:CAA49220.1; PID:g15416
R:Kallman, A.V.; Kiyukov, V.M.; Bayev, A.A.
Nucleic Acids Res. 16, 6230, 1988
A:Title: The nucleotide sequence of bacteriophage T5 DNA at the region between early
A:Reference number: S01982; MUID:88289370
A:Accession: S01984
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 934-985, 'A', 987-1396 <KA2>
A:Cross-references: EMBL:X07559
C:Genetics:
A:Gene: ltf
C:Keywords: late protein; tail fiber

Query Match 1.8%; Score 8; DB 2; Length 1396;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 QSATQAS 61
|||||
DB 87 QSATQAS 94

RESULT 6
S61050

Probable membrane protein YDL163w - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: hypothetical protein D1505
C:Species: *Saccharomyces cerevisiae*
C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 20-Jun-2000
R:Poehl, T.M.
submitted to the EMBL Data Library, November 1995
A:Reference number: S61010
A:Accession: S61050
A:Molecule type: DNA
A:Residues: 1-100 <POH>
A:Cross-references: EMBL:Z67750; NID:g1061256; PID:g1061276
R:Poehl, T.M.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67708
A:Accession: S67715
A:Molecule type: DNA
A:Residues: 1-100 <POW>
A:Cross-references: EMBL:Z74212; NID:g1431258; PID:g1431260; MIPS:YDL163w
A:Experimental source: strain S288C
C:Genetics:
A:Map position: 4L
C:Superfamily: Saccharomyces probable membrane protein YDL163w
C:Keywords: transmembrane protein
F:3-19/Domain: transmembrane #status predicted <TM>

Query Match 1.6%; Score 7; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 LTLSTLS 14
|||||
DB 26 LTLSTLS 32


```
RESULT 7
F86676
hypothetical protein ptca [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 23-Mar-2001
C:Accession: F86676
R:Boletín, A.; Mincker, P.; Manger, S.; Jallón, O.; Malarme, K.; Weissensbach, J.; Ehrlich
Genome Res. in press, 2001
A:Title: The complete genome sequence of the lactic acid bacterium.
A:Reference number: A86625
A:Accession: F86676
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-116 <STO>
A:Cross-references: GB:AE005176; NID:912723289; PIDN:AAK04512.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: ptca

Query Match 1.6%; Score 7; DB 2; Length 116;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 AIOQAKA 26
|||||
Db 34 AIOQAKA 40

RESULT 8
E75377
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: E75377
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: E75377
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-122 <WHI>
A:Cross-references: GB:AE002002; GB:AE000513; NID:96459345; PIDN:AAF11150.1; PID:9645935
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1580
A:Map position: 1

Query Match 1.6%; Score 7; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 EHEPELY 76
|||||
Db 115 EHEPELY 121

RESULT 9
E86498
pts 11A protein [imported] - Chlamydia pneumoniae (strain J138)
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: E86498
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349
A:Accession: E86498
```

```
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-158 <STO>
A:Cross-references: GB:BA000008; NID:98978434; PIDN:BA98271.1; GSPDB:GN00142
A:Experimental source: strain J138
C:Genetics:
A:Gene: ptsN1
C:Superfamily: phosphotransferase system enzyme II; phosphotransferase system mannito

Query Match 1.6%; Score 7; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 LSTLTLS 11
|||||
Db 126 LSTLTLS 132

RESULT 10
E72124
pts 11a protein - Chlamydia pneumoniae (strain CWL029)
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
C:Accession: E72124
R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606
A:Accession: E72124
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-158 <ARN>
A:Cross-references: GB:AE001591; GB:AE001363; NID:94376311; PIDN:AAD18213.1; PID:9437
A:Experimental source: strain CWL029
C:Genetics:
A:Gene: ptsN1
C:Superfamily: phosphotransferase system enzyme II; phosphotransferase system mannito
F:28-156/Domain: phosphotransferase system mannitol-specific enzyme II factor III hom

Query Match 1.6%; Score 7; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 LSTLTLS 11
|||||
Db 126 LSTLTLS 132

RESULT 11
A54506
tubulin alpha-1 chain - Plasmodium yoelii (fragment)
C:Species: Plasmodium yoelii
C:Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 13-Aug-1999
C:Accession: A54506
R:Akella, R.; Arasu, P.; Vaidya, A.B.
Mol. Biochem. Parasitol. 30, 165-174, 1988
A:Title: Molecular clones of alpha-tubulin genes of Plasmodium yoelii reveal an unusu
A:Reference number: A54506; MUID:89014607
A:Accession: A54506
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-197 <AKE>
A:Cross-references: GB:M29816; NID:9160729; PIDN:AAA29779.1; PID:9160730
C:Superfamily: tubulin

Query Match 1.6%; Score 7; DB 2; Length 197;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 ITKRSI 120
```

Db 82 IKTKRSI 88

RESULT 12

TI9438
hypothetical protein C25A1.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: TI9438

R:Morimore, B.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19124

A:Accession: TI9438

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-219 <NTL>

A:Cross-references: EMBL:281038; PIDN:CAB02762.1; GSPDB:GN00019; CESP:C25A1.1

A:Experimental source: clone C25A1

C:Genetics:

A:Gene: CESP:C25A1.1

A:Map position: 1

A:Introns: 33/1; 66/1; 151/2

Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 219;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 158 KDAKOLE 164

Db 52 KDAKOLE 58

RESULT 13

C84976

tRNA (guanine-N1)-methyltransferase (EC 2.1.1.31) [imported] - Buchnera sp. (strain APS

C:Species: Buchnera sp.

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001

C:Accession: C84976

R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.

Nature 407, 81-86, 2000

A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A

A:Reference number: A84930; MUID:20445173

A:Accession: C84976

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-237 <STO>

A:Cross-references: GB:AP000398; GSPDB:GN00144

A:Experimental source: strain APS

C:Genetics:

A:Gene: trmD; BU396

C:Superfamily: tRNA (guanine-N1) methyltransferase

C:Keywords: methyltransferase; S-adenosylmethionine

Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 237;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 LAIQOAK 25

Db 78 LAIQOAK 84

RESULT 14

A61544

tubulin alpha chain - common octopus (fragment)

C:Species: Octopus vulgaris (common octopus)

C:Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 10-Jul-1998

C:Accession: A61544

R:Zinov'eva, R.D.; Alelnikova, K.S.; Tomarev, S.I.

Dokl. Akad. Nauk SSSR 302, 462-467, 1988

A:Title: Isolation and structural characterization of cDNAs coding for alpha-tubulin

A:Reference number: A61544

A:Accession: A61544

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-240 <ZIN>

C:Superfamily: tubulin

F:239-240/Cleavage site: Glu-Tyr (tubulin-specific carboxypeptidase) #status predicte

F:239-240/Cross-link: peptide (Glu-Tyr) (by tubulin-tyrosine ligase) #status predicte

Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 240;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTKRSI 120

Db 124 IKTKRSI 130

RESULT 15

A44959

coat protein - potato virus Y (strain N) (fragment)

C:Species: potato virus Y, PVY

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Nov-2000

C:Accession: A44959

R:Hay, J.M.; Fellowes, A.P.; Timmerman, G.M.

Arch. Virol. 107, 111-122, 1989

A:Title: Nucleotide sequence of the coat protein gene of a necrotic strain of potato

A:Reference number: A44959; MUID:90025790

A:Accession: A44959

A:Status: preliminary

A:Molecule type: genomic RNA

A:Residues: 1-264 <HAY>

A:Cross-references: GB:M22470; NID:g333508; PIDN:AAA47182.1; PID:g333509

C:Superfamily: tobacco etch virus genome polypeptide

Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 264;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 156 TKKDAKQ 162

Db 8 TKKDAKQ 14

Search completed: May 9, 2002, 12:56:34
Job time: 604 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2002, 10:23:31; Search time 71.76 Seconds

(without alignments)
456.249 Million cell updates/sec

Title: US-09-787-083-8

Perfect score: 442
Sequence: 1 MKVSLSTLTLSTLPCFALLA.....YHNEATSPGVGLMNDMGL 442

Scoring table:

OLIGO
Gapop 60.0, Gapext 60.0

Searched: 522463 seqs, 74073290 residues

Word size: 15

Total number of hits satisfying chosen parameters: 4

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

A_Geneseq_1101:*

1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	442	100.0	442	21	AAV85271
2	327	74.0	442	21	BAVB034 amino acid
3	316	71.5	442	21	BAVB034 amino acid
4	289	65.4	442	21	AAV85269
					BAVB034 amino acid

ALIGNMENTS

RESULT 1
AAV85271

ID	AAV85271 standard; Protein: 442 AA.
XX	AC AAV85271;
XX	29-JUN-2000 (first entry)
XX	DE BAVB034 amino acid sequence #4.
XX	Moraxella catarrhalis infection; BASB034: diagnosis; staging;
KW	sinusitis; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW	sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW	hearing loss; antibacterial drug.
XX	Moraxella catarrhalis.
XX	MO200015802-A1.
PN	23-MAR-2000.
XX	14-SEP-1999; 99WO-EP06781.
PR	14-SEP-1998; 98GB-0020002.
XX	(SMIK) SWITHKLINE BEECHAM BIOLOGICALS.
PI	Ruelle J;
XX	WPI; 2000-271440/23.
DR	N-PSDB: AAA10703.
XX	Novel BASB034 polynucleotides and polypeptides from Moraxella
PT	catarrhalis used to prepare vaccines against bacterial infections
PS	Claim 3; Page 69; 106pp; English.
XX	This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC	strain Mc2969. The invention relates to BASB034 polypeptides from
CC	M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
CC	polynucleotides and polypeptides may be employed as research reagents and
CC	material for the discovery of treatments and diagnostics for diseases,
CC	particularly human diseases. They are particularly used to diagnose and
CC	treat M. catarrhalis infections. They can be used for diagnosis of
CC	disease, staging of disease, or determining response of an infectious
CC	organism to drugs. The polynucleotides may be used as a source for
CC	hybridization probes, and for screening of genetic mutations, serotype,
CC	organism or strain identification, identification of mutations in BASB034
CC	sequences, and as components of arrays which are useful for diagnostic
CC	and prognostic purposes. The polypeptides can be used to produce
CC	antibodies. The polypeptides can also be used in vaccine formulations,
CC	and to identify agonists and antagonists. The polypeptides, antibodies,
CC	agonists and antagonists (which are bacteriostatic) are used for the
CC	treatment and prevention of diseases such as otitis media in infants and
CC	children, pneumonia in elderly, sinusitis, nosocomial infections and
CC	invasive diseases, and chronic otitis media with hearing loss. The
CC	polypeptides, agonists and antagonists are also used for screening of
CC	antibacterial drugs. The BASB034 products of the invention can be used
CC	screen for new antibacterial compounds that may target resistant
CC	bacteria.
XX	
XX	
SQ	Sequence 442 AA:
QY	Query Match 100.0%; Score 442; DB 21; Length 442;
DB	Best Local Similarity 100.0%; Pred. No. 0;
DB	Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MKVSLSTLTLSTLPCFALLAQAAPVPAFVVEVSENDIGDNLPLDVGSAQSA 60
DB	1 MKVSLSTLTLSTLPCFALLAQAAPVPAFVVEVSENDIGDNLPLDVGSAQSA 60
QY	61 STDTPANPLDEHEPELTYTALENKTMKINCSALNODIMRLACYDTLVHGETPAVITKRSI 120
DB	61 STDTPANPLDEHEPELTYTALENKTMKINCSALNODIMRLACYDTLVHGETPAVITKRSI 120

PD 23-MAR-2000.
 XX 14-SEP-1999; 99WO-EP06781.
 PF 14-SEP-1998; 98GB-0020002.
 PR (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 PA Ruelle J;
 XX WPI: 2000-271440/23.
 DR N-PSDB; AAA10702.
 XX
 PT Novel BASB034 polynucleotides and polypeptides from Moraxella
 catarrhalis used to prepare vaccines against bacterial infections -
 PS Claim 3; Page 68; 106pp; English.

CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 CC strain Mc2913. The invention relates to BASB034 polypeptides from
 CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
 CC polynucleotides and polypeptides may be employed as research reagents and
 CC material for the discovery of treatments and diagnostics for diseases,
 CC particularly human diseases. They are particularly used to diagnose and
 CC treat M. catarrhalis infections. They can be used for diagnosis of
 CC disease, staging of disease, or determining response of an infectious
 CC organism to drugs. The polynucleotides may be used as a source for
 CC hybridization probes, and for screening of genetic mutations, serotype,
 CC organism or strain identification, identification of mutations in BASB034
 CC sequences, and as components of arrays which are useful for diagnostic
 CC and prognostic purposes. The polypeptides can be used to produce
 CC antibodies. The polypeptides can also be used in vaccine formulations,
 CC and to identify agonists and antagonists. The polypeptides, antibodies,
 CC agonists and antagonists (which are bacteriostatic) are used for the
 CC treatment and prevention of diseases such as otitis media in infants and
 CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
 CC invasive diseases, and chronic otitis media with hearing loss. The
 CC polypeptides, agonists and antagonists are also used for screening of
 CC antibacterial drugs. The BASB034 products of the invention can be used
 CC screen for new antibacterial compounds that may target resistant
 CC bacteria.
 CC
 XX
 XX
 SO Sequence 442 AA;

Query Match 71.5%; Score 316; DB 21; Length 442;
 Best Local Similarity 99.8%; Pred. No. 1.3e-298;
 Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 26 AVPRPVAFVDEVRSENDLGDNEPRIDVQSNQTSASTDTANPLDEHEPELYTTALENKTM 85
 DB 26 avprpva fvevrse ndlgdne pridvqsnqtsastdtanpldehepe lyttalenktm 85
 QY 86 LINSALNODIMRLACYDTLVHGETPAVITKRSIRLDETITMOTIKGPOVVYQETDPPI 145
 DB 86 linsalng dlmrlacyd tlvhgetpav itkrsir ldetitmo tikgpo vvyyqetd ppi 145
 QY 146 FLMGNEKGMETTKDAKOLETAAKQFTPLUSFDLDRNNTPLMSSRPNNMPTVLDIFMHGK 205
 DB 146 flmgnek gm ettkda koleyaa kqftpl usfdldr nn tplmssr pnnmptv ldifmhgk 205
 QY 206 PNRSPNPSHEAKQFTNEPAPLKFQVSVKAAEDLMDTSDIDMFGYQOSHMQIFN 265
 DB 206 pnrspnps sheakqft ne papl kfqvsv kaaedl mdt sdi dmfgyq oshmqifn 265
 QY 266 GKNSRPRFVHDYOPFELFTQPVVYSDLPWDGKVRMIGMGAHHSNGESAKLSRSWNRAYLM 325
 DB 266 gknsrpr fvdhyop felftqpv vy sdlp wd gkvr migmga hhsnges akslrs wnraylm 325
 QY 326 AGMEKMLTYWPRITWRIKFGSGSPDDNDLIDYGYGDFRFLYOLENKSNTSGTVRY 385
 DB 326 agmekml tywpr itwri kfgsgs pddnd lidygy gdf rfl yolenks ntsgtv ry 385

OY 386 NPSRGKALODDYVYPIGKGISGYSFOYFGYGSGLIDYNHEATSPFGYGLMLNDMWGL 442
 DB 386 npsrgkal odyvy p i gk gisg ysfo yfgygs glidyn heatsp fgygl mldmwgl 442

RESULT 4
 ID AAY85269 standard; Protein: 442 AA.
 AC AAY85269;
 XX
 DT 29-JUN-2000 (first entry)
 XX

DE BASB034 amino acid sequence #2.

XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;
 KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
 KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
 KW hearing loss; antibacterial drug.

XX Moraxella catarrhalis.

OS WO200015802-A1.

PN 23-MAR-2000.

PF 14-SEP-1999; 99WO-EP06781.

PR 14-SEP-1998; 98GB-0020002.

PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

PI Ruelle J;

DR WPI: 2000-271440/23.

DR N-PSDB; AAA10701.

PT Novel BASB034 polynucleotides and polypeptides from Moraxella

catarrhalis used to prepare vaccines against bacterial infections -

PS Claim 3; Page 67; 106pp; English.

CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 CC strain Mc2908. The invention relates to BASB034 polypeptides from
 CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
 CC polynucleotides and polypeptides may be employed as research reagents and
 CC material for the discovery of treatments and diagnostics for diseases,
 CC particularly human diseases. They are particularly used to diagnose and
 CC treat M. catarrhalis infections. They can be used for diagnosis of
 CC disease, staging of disease, or determining response of an infectious
 CC organism to drugs. The polynucleotides may be used as a source for
 CC hybridization probes, and for screening of genetic mutations, serotype,
 CC organism or strain identification, identification of mutations in BASB034
 CC sequences, and as components of arrays which are useful for diagnostic
 CC and prognostic purposes. The polypeptides can be used to produce
 CC antibodies. The polypeptides can also be used in vaccine formulations,
 CC and to identify agonists and antagonists. The polypeptides, antibodies,
 CC agonists and antagonists (which are bacteriostatic) are used for the
 CC treatment and prevention of diseases such as otitis media in infants and
 CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
 CC invasive diseases, and chronic otitis media with hearing loss. The
 CC polypeptides, agonists and antagonists are also used for screening of
 CC antibacterial drugs. The BASB034 products of the invention can be used
 CC screen for new antibacterial compounds that may target resistant
 CC bacteria.
 CC
 XX
 XX
 SO Sequence 442 AA;

Query Match 65.4%; Score 289; DB 21; Length 442;
 Best Local Similarity 99.7%; Pred. No. 2.3e-272;
 Matches 389; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 53 VQATQASASTDTANPLDEHEPELYTTALENKMLINCASALNDIMRLACYDTLVHGETPA 112
|||||
Db 53 vgsatqsaastcltanpldehepeylttalenkmlincsalngdlmracydtlvhgetpa 112
QY 113 VIKTKRSIRLDETIMQTIKGPQVVOETDPFELMGNEKGMLTKKDAKOLEYAAKOFTP 172
|||||
Db 113 viktksirldetimgtlikgkpvvygetcdplflmgnekgmllckdakqleyaakqftcp 172
QY 173 LSLSPDLDRNTPPLWSSRPHPMYVLPIFMHGKPNRSPNTPSHEAKOFTPEFRAPELKP 232
|||||
Db 173 lsisfldrnpplwsrphnmyvlplfmhgpnrspntpshearqfipnefrapelkf 232
QY 233 QVSVKYKAAEDLMGTSDLMFGITQOSHWCIFNGKNSRPRVHDYQPEIFLTQPVYSDLP 292
|||||
Db 233 qvsvkkaaedlwgtsdliwfytlqgshwqlfngknsrprvhdypeliftqpyvsdlp 292
QY 293 WDGKVRMIGMGAVHNSNGESAKLSRSWNRAYLMAGMEWKNLTVMPRIMGRIFKESGSOP 352
|||||
Db 293 wdgkvrmlgmavhnsngesaklsrswnraylmagmewknlvmprlmgrlfksgsgsp 352
QY 353 DDNPDLIDYVYGDVRFELYOLENKSNIQTVRYNPRSGKALQUDYVYPLGKGISGYFOI 412
|||||
Db 353 ddnpdlldyvygdvrflyglenknsisgtvrynprrsgkalqldyvyplgkglsgyfoi 412
QY 413 FQGTGOSLIDYNNHEATSFQVGLMLNDMGL 442
|||||
Db 413 fggysgslidyneatsfyglmndmgml 442

Search completed: May 3, 2002, 10:23:32
Job time: 253 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2002, 10:19:19 ; Search time 71.76 Seconds
(without alignments)
456.249 Million cell updates/sec

Title: US-09-787-083-2

Perfect score: 442
Sequence: 1 MKVSLSTLTLSTLSCFALIA.....YHHEATSFVGLMLNDMGL 442

Scoring table:
OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 522463 seqs, 74073290 residues

Word size : 15

Total number of hits satisfying chosen parameters: 4

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :
1: A_Geneseq_1101.*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
23: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	442	100.0	442	21	AAV85268
2	327	74.0	442	21	AAV85271
3	224	50.7	442	21	AAV85269
4	224	50.7	442	21	AAV85270

ALIGNMENTS

RESULT 1
AAV85268

ID	AAV85268	standard; protein: 442 AA.
AC	AAV85268;	
XX	29-JUN-2000	(first entry)
DE	BASB034	amino acid sequence #1.
KW	Moraxella catarrhalis	infection; BASB034; diagnosis; staging; vaccine; bacteriostatic; treatment; otitis media; pneumonia; sinusitis; nosocomial infection; invasive disease; chronic otitis media; hearing loss; antibacterial drug.
OS	Moraxella	catarrhalis.
XX	MO200015802-A1.	
XX	23-MAR-2000.	
XX	14-SEP-1999;	99WO-EP06781.
XX	14-SEP-1998;	98GB-0020002.
XX	(SMIK)	SMITHKLINE BEECHAM BIOLOGICALS.
XX	Ruelle J;	
XX	WPI; 2000-271440/23.	
DR	N-PSDB; AAA10700.	
XX	Novel BASB034	polynucleotides and polypeptides from Moraxella catarrhalis used to prepare vaccines against bacterial infections
PT	Claim 3; Fig 2; 106pp; English.	
XX	This sequence represents a Moraxella catarrhalis BASB034 polypeptide from strain Mc2931 (ATCC 43617). The invention relates to BASB034 polypeptides from M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034 polynucleotides and polypeptides may be employed as research reagents and material for the discovery of treatments and diagnostics for diseases, particularly human diseases. They are particularly used to diagnose and treat M. catarrhalis infections. They can be used for diagnosis of disease, staging of disease, or determining response of an infectious organism to drugs. The polynucleotides may be used as a source for hybridization probes, and for screening of genetic mutations, serotype, organism or strain identification, identification of mutations in BASB034 sequences, and as components of arrays which are useful for diagnostic and prognostic purposes. The polypeptides can be used to produce antibodies. The polypeptides can also be used in vaccine formulations, and to identify agonists and antagonists. The polypeptides, antibodies, agonists and antagonists (which are bacteriostatic) are used for the treatment and prevention of diseases such as otitis media in infants and children, pneumonia in elderly, sinusitis, nosocomial infections and invasive diseases, and chronic otitis media with hearing loss. The polypeptides, agonists and antagonists are also used for screening of antibacterial drugs. The BASB034 products of the invention can be used screen for new antibacterial compounds that may target resistant bacteria.	
SO	Sequence	442 AA;
QY	Query Match	100.0%; Score 442; DB 21; Length 442;
DB	Best Local Similarity	100.0%; Pred. No. 0;
DB	Matches	442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1	MKVSLSTLTLSTLSCFALIAQAAPVAFVDEVSNDIGDNEPLIDVOSATOSA 60
DB	1	mkvslstltlslscfaliaaigagavpnpvafvdevsndigqdnelpidvgaaqsa 60
QY	61	STDTANPLDEHEPELYTALNKMTLINCASLNDIMRLACYDPLVHGEPVAVIKTKRSI 120
DB	61	stdtanpldehepeelytalcnkmlncasalndimglmlacydplvhetpaviktkrsi 120

PD 23-MAR-2000.
XX 14-SEP-1999; 99WO-EP06781.
PF 14-SEP-1998; 98GB-0020002.
PR 14-SEP-1998; 98GB-0020002.
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
PA
XX
XX Ruelle J;
PI
XX WPI: 2000-271440/23.
DR N-PSDB; AAA10701.
XX
XX Novel BASB034 polynucleotides and polypeptides from Moraxella
PT catarrhalis used to prepare vaccines against bacterial infections
XX
XX Claim 3; Page 67; 106pp; English.
XX
XX This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC strain Mc2908. The invention relates to BASB034 polypeptides from
CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
CC polynucleotides and polypeptides may be employed as research reagents and
CC material for the discovery of treatments and diagnostics for diseases,
CC particularly human diseases. They are particularly used to diagnose and
CC treat M. catarrhalis infections. They can be used for diagnosis of
CC disease, staging of disease, or determining response of an infectious
CC organism to drugs. The polynucleotides may be used as a source for
CC hybridization probes, and for screening of genetic mutations, serotype,
CC organism or strain identification, identification of mutations in BASB034
CC sequences, and as components of arrays which are useful for diagnostic
CC and prognostic purposes. The polypeptides can be used to produce
CC antibodies. The polypeptides can also be used in vaccine formulations,
CC and to identify agonists and antagonists. The polypeptides, antibodies,
CC agonists and antagonists (which are bacteriostatic) are used for the
CC treatment and prevention of diseases such as otitis media in infants and
CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
CC invasive diseases, and chronic otitis media with hearing loss. The
CC polypeptides, agonists and antagonists are also used for screening of
CC antibacterial drugs. The BASB034 products of the invention can be used
CC screen for new antibacterial compounds that may target resistant
CC bacteria.
CC
XX
XX Sequence 442 AA:
SQ
Query Match 50.7%; Score 224; DB 21; Length 442;
Best Local Similarity 100.0%; Pred. No. 9.2e-209;
Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 219 QFTFNERAPRLKQVSVKKAADLMGTSDDLMEFGYTQOSHMQIFNGKNSRRPRVNDYQ 278
DB 219 qftfnetrapelkqvsvkkaaedlmgtsddlwfgytqshmqifngknsrprvndyq 278
QY 279 PEILTPQVYSDDLPMWCKVRMIGWAVHNSGESAKLSRSNRRAYLMAAGMEKMLTYMPR 338
DB 279 peiltfpvysdplmwckvrmigmavhnsngesaklsrsnrraylماغمءكلتءمءر 338
QY 339 IWGRIFKESGSGOPDDNDPILDYGYGVRFYOLENKSNTSGTVRYNPRSGKALOLDY 398
DB 339 iwgrlfksgsgopddndpildygygvrfyolenksntsgtvrynpءsgkاللدء 398
QY 399 VYPIRGKISGFOFEOGYGOSLIDYNHEATSEFGGLMLNDMMGL 442
DB 399 vyprlgkisgfofeygygoslidynheatsfgvglmndmmgl 442
RESULT 4
ID AAY85270 standard; Protein; 442 AA.
XX
XX AAY85270;
XX
XX 29-JUN-2000 (first entry)

XX
XX BASB034 amino acid sequence #3.
DE
XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;
XX vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
XX sinusitis; nosocomial infection; invasive disease; chronic otitis media;
XX hearing loss; antibacterial drug.
XX
XX Moraxella catarrhalis.
XX
XX WO200015802-A1.
XX
XX 23-MAR-2000.
XX
XX 14-SEP-1999; 99WO-EP06781.
XX
XX 14-SEP-1998; 98GB-0020002.
XX
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Ruelle J;
XX
XX WPI: 2000-271440/23.
XX
XX N-PSDB; AAA10702.
XX
XX Novel BASB034 polynucleotides and polypeptides from Moraxella
PT catarrhalis used to prepare vaccines against bacterial infections
XX
XX Claim 3; Page 68; 106pp; English.
XX
XX This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC strain Mc2913. The invention relates to BASB034 polypeptides from
CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
CC polynucleotides and polypeptides may be employed as research reagents and
CC material for the discovery of treatments and diagnostics for diseases,
CC particularly human diseases. They are particularly used to diagnose and
CC treat M. catarrhalis infections. They can be used for diagnosis of
CC disease, staging of disease, or determining response of an infectious
CC organism to drugs. The polynucleotides may be used as a source for
CC hybridization probes, and for screening of genetic mutations, serotype,
CC organism or strain identification, identification of mutations in BASB034
CC sequences, and as components of arrays which are useful for diagnostic
CC and prognostic purposes. The polypeptides can be used to produce
CC antibodies. The polypeptides can also be used in vaccine formulations,
CC and to identify agonists and antagonists. The polypeptides, antibodies,
CC agonists and antagonists (which are bacteriostatic) are used for the
CC treatment and prevention of diseases such as otitis media in infants and
CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
CC invasive diseases, and chronic otitis media with hearing loss. The
CC polypeptides, agonists and antagonists are also used for screening of
CC antibacterial drugs. The BASB034 products of the invention can be used
CC screen for new antibacterial compounds that may target resistant
CC bacteria.
CC
XX
XX Sequence 442 AA:
SQ
Query Match 50.7%; Score 224; DB 21; Length 442;
Best Local Similarity 100.0%; Pred. No. 9.2e-209;
Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 219 QFTFNERAPRLKQVSVKKAADLMGTSDDLMEFGYTQOSHMQIFNGKNSRRPRVNDYQ 278
DB 219 qftfnetrapelkqvsvkkaaedlmgtsddlwfgytqshmqifngknsrprvndyq 278
QY 279 PEILTPQVYSDDLPMWCKVRMIGWAVHNSGESAKLSRSNRRAYLMAAGMEKMLTYMPR 338
DB 279 peiltfpvysdplmwckvrmigmavhnsngesaklsrsnrraylماغمءكلتءمءر 338
QY 339 IWGRIFKESGSGOPDDNDPILDYGYGVRFYOLENKSNTSGTVRYNPRSGKALOLDY 398
DB 339 iwgrlfksgsgopddndpildygygvrfyolenksntsgtvrynpءsgkاللدء 398

OY 399 VYPLGKISGYFOIFQGYGSLIDYNHEATSEGVGLMNDMGL 442
|||
Db 399 VYPLGKISGYFq1fq9y9gslidynheatsfgvglmndmg1 442

Search completed: May 3, 2002, 10:23:31
Job time: 252 sec

PR 14-SEP-1998; 98GB-0020002.
XX
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Ruelle J;
XX
XX WPI: 2000-2714A0/23.
DR N-PSDB; AAA10703.
XX
PT Novel BASB034 polynucleotides and polypeptides from Moraxella
PT catarrhalis used to prepare vaccines against bacterial infections
XX
XX Claim 3; Page 69; 106pp; English.
XX
XX This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC strain Mc2969. The invention relates to BASB034 polypeptides from
CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
CC polynucleotides and polypeptides may be employed as research reagents and
CC material for the discovery of treatments and diagnostics for diseases,
CC particularly human diseases. They are particularly used to diagnose and
CC treat M. catarrhalis infections. They can be used for diagnosis of
CC disease, staging of disease, or determining response of an infectious
CC organism to drugs. The polynucleotides may be used as a source for
CC hybridization probes, and for screening of genetic mutations, serotype,
CC organism or strain identification, identification of mutations in BASB034
CC sequences, and as components of arrays which are useful for diagnostic
CC and prognostic purposes. The polypeptides can be used to produce
CC antibodies. The polypeptides can also be used in vaccine formulations,
CC and to identify agonists and antagonists. The polypeptides, antibodies,
CC agonists and antagonists (which are bacteriostatic) are used for the
CC treatment and prevention of diseases such as otitis media in infants and
CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
CC invasive diseases, and chronic otitis media with hearing loss. The
CC polypeptides, agonists and antagonists are also used for screening of
CC antibacterial drugs. The BASB034 products of the invention can be used
CC screen for new antibacterial compounds that may target resistant
CC bacteria.
XX
XX Sequence 442 AA:
SQ

Query Match 65.4%; Score 289; DB 21; Length 442;
Best Local Similarity 99.7%; Pred. No. 7.6e-273;
Matches 389; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 53 VQASQTASDTGANPLDEHEBELYTALLENKTMILNCSALNODIMRLACYDTLVHGETPA 112
DB 53 vgsatgsastltanpldehebelyltalenktmlncsalngdmlrlacydtllvhgetpa 112

QY 113 VIKRKRSRLDELTWQITKGRPOVYVQETDPIFLMGNEKMLTKKAKOLEYAAKQFTP 172
DB 113 vikrkrsrldeletwqitkgrpovyvqetdprilfmgnekmllkkakqleyaakqftt 172

QY 173 LLSLEFDDRRNTPLMSSRNPNRYVLPPIFMHGKPNRSPNPSHAROPTPEFAPELKTF 232
DB 173 llslefddrrntplmssrnpnrpnyvlprifmhkgpnrsnpnsharoptpefapekft 232

QY 233 QVSVKVAABEDLMGTSDDLWFGYTGQSHWOIFNGKNSRPFVRHDYQBEIFLPQVYSDLP 292
DB 233 qvsavkvaabedlmgtsddlwfgytgqshwqifngknsrpfvrhdyqbeifltpqvysdlp 292

QY 293 WDGVVRVRIQMGAVVHNSGESAKLSRSWNRATLTMAGMEKMLTVNPRIMGRIFKFGSGSQP 352
DB 293 wdgvvrvmgmavvhnsgesaklsrswnratlmagmekmltvnprimgri fkgsgsqp 352

QY 353 DDNDPDLIDYGYGVDFVLEENKSNISGTVRPNRSGKALODYVYPLGKGISGYFOI 412
DB 353 ddndpdlidygygvdfvleenksnisgtrvprnsgkalgldyvyplgkgisgyfqi 412

QY 413 FQGYGQSLIDYNNHEATSFVGVLMDNMWGL 442
DB 413 fggygqslidyynheatsfgvgvlmdnmwgl 442

RESULT 4
AAV85268
ID AAV85268 standard; Protein; 442 AA.
XX
XX AAV85268;
AC
XX 29-JUN-2000 (first entry)
DT
XX
XX BASB034 amino acid sequence #1.
DE
XX
XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW hearing loss; antibacterial drug.
XX
XX Moraxella catarrhalis.
OS
XX
XX W0200015802-A1.
PN
XX
XX 23-MAR-2000.
PD
XX
XX 14-SEP-1999; 99WO-EP06781.
PF
XX
XX 14-SEP-1998; 98GB-0020002.
PR
XX
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
PA
XX
XX Ruelle J;
PI
XX
XX WPI: 2000-2714A0/23.
DR N-PSDB; AAA10700.
DR
XX
XX Novel BASB034 polynucleotides and polypeptides from Moraxella
PT catarrhalis used to prepare vaccines against bacterial infections
XX
XX Claim 3; Fig 2; 106pp; English.
XX
XX This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC strain Mc2931 (ATCC 43617). The invention relates to BASB034 polypeptides
CC from M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The
CC BASB034 polynucleotides and polypeptides may be employed as research
CC reagents and material for the discovery of treatments and diagnostics for
CC diseases, particularly human diseases. They are particularly used to
CC diagnose and treat M. catarrhalis infections. They can be used for
CC diagnosis of disease, staging of disease, or determining response of an
CC infectious organism to drugs. The polynucleotides may be used as a source
CC for hybridization probes, and for screening of genetic mutations,
CC serotype, organism or strain identification, identification of mutations
CC in BASB034 sequences, and as components of arrays which are useful for
CC diagnostic and prognostic purposes. The polypeptides can be used to
CC produce antibodies. The polypeptides can also be used in vaccine
CC formulations, and to identify agonists and antagonists. The polypeptides,
CC antibodies, agonists and antagonists (which are bacteriostatic) are used
CC for the treatment and prevention of diseases such as otitis media in
CC infants and children, pneumonia in elderlies, sinusitis, nosocomial
CC infections and invasive diseases, and chronic otitis media with hearing
CC loss. The polypeptides, agonists and antagonists are also used for
CC screening of antibacterial drugs. The BASB034 products of the invention
CC can be used screen for new antibacterial compounds that may target
CC resistant bacteria.
XX
XX Sequence 442 AA:
SQ

Query Match 50.7%; Score 224; DB 21; Length 442;
Best Local Similarity 100.0%; Pred. No. 1.4e-209;
Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 219 QTFNFRAPELKQVSVKVAABEDLMGTSDDLWFGYTGQSHWOIFNGKNSRPFVRHDYQ 278
DB 219 qtfpnefrapelkfqsvkvaabedlmgtsddlwfgytgqshwqifngknsrpfvrhdyq 278

QY 279 PEIFLTOPVYSDLPWOCKVBMIGAVHSHNGESAKLSRSGWNRAYLMAGHEMKNLTVMPR 338
| | | | |
Db 279 pelfltpvysdipwdgkvrmigmavhshngesaklsrswrnylmagmewknlvmpc 338
| | | | |
QY 339 IMGRIFKEGSGSQPDDNPDILDYRGDVRFLYOLENKSNSIGTVRYRNPBSGKALQLDY 398
| | | | |
Db 339 lwgrlftkgs9sqddnplldyygydvrllylenksnlsgrlvrynpbsgk9alqldy 398
| | | | |
QY 399 VYPLGKGISGYFOIFQGYGQSLIDYNHEATSFGVGLMLNDMGL 442
| | | | |
Db 399 vylpgkgsigylfglfggygslldynheatsfgvglmndmgl 442
| | | | |

Search completed: May 3, 2002, 10:23:31
Job time: 252 sec

QY 121 RLDETIWQTIKGRQVYVOETTDPIFLMGNEKGLTKKAKOLEYAAKQFTPLSLSPDLL 180
| | | | |
DB 121 RLDETIWQTIKGRQVYVOETTDPIFLMGNEKGLTKKAKOLEYAAKQFTPLSLSPDLL 180
QY 181 RNNTPMLSSPPHNPMMYLPFIEMHGKPNRSPNTPSHEAROPFNFRAPDELKFOQSVYVKA 240
| | | | |
DB 181 RNNTPMLSSPPHNPMMYLPFIEMHGKPNRSPNTPSHEAROPFNFRAPDELKFOQSVYVKA 240
QY 241 AEDLMGTDSDLMFGYTOOSHMOIFNGKNSRPFRRVNDYQPEIFLTPQVYSDLPMDMGKFRMI 300
| | | | |
DB 241 AEDLMGTDSDLMFGYTOOSHMOIFNGKNSRPFRRVNDYQPEIFLTPQVYSDLPMDMGKFRMI 300
QY 301 GKGAVHNSGESAKLSRSMNRATYLMAGMEKMLTVPFRIGRIFKESSGSDPDNDPILD 360
| | | | |
DB 301 GKGAVHNSGESAKLSRSMNRATYLMAGMEKMLTVPFRIGRIFKESSGSDPDNDPILD 360
QY 361 YVGGDVRFLYOLENKSNTSGTVYRNSRSGKALOLDYVYPLGCTISGYFOIFQVYQOSL 420
| | | | |
DB 361 YVGGDVRFLYOLENKSNTSGTVYRNSRSGKALOLDYVYPLGCTISGYFOIFQVYQOSL 420
QY 421 IDYNHEATSEFGVGLMDNMGL 442
| | | | |
DB 421 IDYNHEATSEFGVGLMDNMGL 442

RESULT 2
AAV85269 standard; Protein; 442 AA.
XX
AC AAV85269:
XX
DT 29-JUN-2000 (first entry)
XX
DE BASB034 amino acid sequence #2.
XX
XX Moraxella catarrahals infection: BASB034; diagnosis: staging;
KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW hearing loss; antibacterial drug.
XX
OS Moraxella catarrahals.
XX
PN W0200015802-A1.
XX
PD 23-MAR-2000.
XX
PF 14-SEP-1999; 99MO-EP06781.
XX
PR 14-SEP-1998; 98GB-0020002.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Ruelle J;
XX
PI WPI: 2000-2714A0/23.
DR N-PDB; AAA10701.
XX
XX Novel BASB034 polynucleotides and polypeptides from Moraxella
PT catarrahals used to prepare vaccines against bacterial infections
XX
PS Claim 3; Page 67; 106pp; English.

This sequence represents a Moraxella catarrahals BASB034 polypeptide from
CC strain Mc2908. The invention relates to BASB034 polypeptides from
CC M. catarrahals strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
CC polynucleotides and polypeptides may be employed as research reagents and
CC material for the discovery of treatments and diagnostics for diseases,
CC particularly human diseases. They are particularly used to diagnose and
CC treat M. catarrahals infections. They can be used for diagnosis of
CC disease, staging of disease, or determining response of an infectious
CC organism to drugs. The polynucleotides may be used as a source for
CC hybridization probes, and for screening of genetic mutations, serotype,

CC organism or strain identification, identification of mutations in BASB034
CC sequences, and as components of arrays which are useful for diagnostic
CC and prognostic purposes. The polypeptides can be used to produce
CC antibodies. The polypeptides can also be used in vaccine formulations,
CC and to identify agonists and antagonists. The polypeptides, antibodies,
CC agonists and antagonists (which are bacteriostatic) are used for the
CC treatment and prevention of diseases such as otitis media in infants and
CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
CC invasive diseases, agonists and chronic otitis media with hearing loss. The
CC polypeptides, agonists and antagonists are also used for screening of
CC antibacterial drugs. The BASB034 products of the invention can be used
CC screen for new antibacterial compounds that may target resistant
CC bacteria.
XX
SQ Sequence 442 AA:
XX

Query Match 88.2%; Score 390; DB 21; Length 442;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 VQSATQASSTDTANPLDEHEPELYTTALENKTMLINCSALNODIMRLACYDTLVHGCTPA 112
| | | | |
DB 53 VQSATQASSTDTANPLDEHEPELYTTALENKTMLINCSALNODIMRLACYDTLVHGCTPA 112
QY 113 VIKTKRSIRLDETIWQTIKGRQVYVOETTDPIFLMGNEKGLTKKAKOLEYAAKQFTPL 172
| | | | |
DB 113 VIKTKRSIRLDETIWQTIKGRQVYVOETTDPIFLMGNEKGLTKKAKOLEYAAKQFTPL 172
QY 173 LSLSEFDLDRNNTPLMSSRPNNPMYLPFIEMHGKPNRSPNTPSHEAROPFNFRAPDELK 232
| | | | |
DB 173 LSLSEFDLDRNNTPLMSSRPNNPMYLPFIEMHGKPNRSPNTPSHEAROPFNFRAPDELK 232
QY 233 QVSVKVAKAEEDLMGTDSDLMFGYTOOSHMOIFNGKNSRPFRRVNDYQPEIFLTPQVYSDLP 292
| | | | |
DB 233 QVSVKVAKAEEDLMGTDSDLMFGYTOOSHMOIFNGKNSRPFRRVNDYQPEIFLTPQVYSDLP 292
QY 293 WDCGVNRIGMGAVHNSGESAKLSRSMNRATYLMAGMEKMLTVPFRIGRIFKESSGSDPD 352
| | | | |
DB 293 WDCGVNRIGMGAVHNSGESAKLSRSMNRATYLMAGMEKMLTVPFRIGRIFKESSGSDPD 352
QY 353 DDNDPILDYVGGDVRFLYOLENKSNTSGTVYRNSRSGKALOLDYVYPLGCTISGYFOI 412
| | | | |
DB 353 DDNDPILDYVGGDVRFLYOLENKSNTSGTVYRNSRSGKALOLDYVYPLGCTISGYFOI 412
QY 413 FQYGGQSLIDYNHEATSEFGVGLMDNMGL 442
| | | | |
DB 413 FQYGGQSLIDYNHEATSEFGVGLMDNMGL 442

RESULT 3
AAV85271 standard; Protein; 442 AA.
XX
AC AAV85271:
XX
DT 29-JUN-2000 (first entry)
XX
DE BASB034 amino acid sequence #4.
XX
XX Moraxella catarrahals infection: BASB034; diagnosis: staging;
KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW hearing loss; antibacterial drug.
XX
OS Moraxella catarrahals.
XX
PN W0200015802-A1.
XX
PD 23-MAR-2000.
XX
PF 14-SEP-1999; 99MO-EP06781..

PR 14-SEP-1998: 98GB-0020002.
XX
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Ruelle J;
PI
XX WPI: 2000-271440/23.
DR N-PSDB; AAA10703.
XX
XX Novel BASB034 polynucleotides and polypeptides from Moraxella
PT catarhalis used to prepare vaccines against bacterial infections
PS Claim 3; Page 69; 106pp: English.
XX
XX This sequence represents a Moraxella catarhalis BASB034 polypeptide from
CC strain Mc2969. The invention relates to BASB034 polypeptides from
CC M. catarhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
CC polynucleotides and polypeptides may be employed as research reagents and
CC material for the discovery of treatments and diagnostics for diseases,
CC particularly human diseases. They are particularly used to diagnose and
CC treat M. catarhalis infections. They can be used for diagnosis of
CC disease, staging of disease, or determining response of an infectious
CC organism to drugs. The polynucleotides may be used as a source for
CC hybridization probes, and for screening of genetic mutations, serotype,
CC organism or strain identification, identification of mutations in BASB034
CC sequences, and as components of arrays which are useful for diagnostic
CC and prognostic purposes. The polypeptides can be used to produce
CC antibodies. The polypeptides can also be used in vaccine formulations,
CC and to identify agonists and antagonists. The polypeptides, antibodies,
CC agonists and antagonists (which are bacteriostatic) are used for the
CC treatment and prevention of diseases such as otitis media in infants and
CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
CC invasive diseases, and chronic otitis media with hearing loss. The
CC polypeptides, agonists and antagonists are also used for screening of
CC antibacterial drugs. The BASB034 products of the invention can be used
CC screen for new antibacterial compounds that may target resistant
CC bacteria.
XX
XX Sequence 442 AA:
SO
Query Match 71.5%; Score 316; DB 21; Length 442;
Best Local Similarity 99.8%; Pred. No. 3.6e-299;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 26 AVPNPVADEVSRSENDLGQDNELPIDVQASATQASTDPANPLDEHEPELYTTALENKTW 85
DB 26 avnpvatdevsrseidvgdnelpidvqatsqastdpanldehepeyltalentkm 85
QY 86 LINCASALNODIMRLACQYTLVHGETPANYIKTRSRIDETTWQTKGAPVYVQETTPDI 145
DB 86 lincsalnqdlmrilacqytlvhgetpaviktrsrirldetlwtqtkgprvyvqettpdi 145
QY 146 FLMSGEKGLTFYKKDAKOLEYAKOFTPLSLSPDLRNNTPMLSSPHPMYVLLPTFMHGK 205
DB 146 flmggekmltkkdaqkleyaakqftplslsfdlrmnplwssrphmvyvllpfmngk 205
QY 206 PNRSPNTPSHEARQFTPNFEFRAPELKFQVSVKAKAEDLWGTDSJLWFGYTQOSHMOJFN 265
DB 206 pnrsptnshearqftpnfeafapelkfqsvvkaaedlwtgdsdlwtfgytqshwqjfn 265
QY 266 GKNSRPFVHYOYPERIFLTQPYSDLPMDGKYRMTGKAGVAHNSGNSAKLSNMRATYIM 325
DB 266 gknsrpfvhydypei fltqpyssdlpmdgkvrmtgkagvahnsgnsaklsrsmratyilm 325
QY 326 AGMEKNLTVMRIRIGRIFFKEGSGSQPDNDPILDYGYGVRFLYQLEKNKSNIGTVRY 385
DB 326 agmeknltvmprlwgri ffrfkegsgsqpddndpildygygvrflyqlenknisgltvry 385
QY 386 NPSGKALQLDYVYPLGKGISGYFOIFQYGOGLIDYNHEATSFVGLMLDMWGL 442
DB 386 npsgkalgldyvyvplgkgisgyfoifqygoqliidynheatsfvglmldmwgl 442

RESULT 4
AAV85268
ID AAV85268 standard; Protein: 442 AA.
XX
XX
AC AAV85268;
XX
DT 29-JUN-2000 (first entry)
XX
DE BASB034 amino acid sequence #1.
XX
KW Moraxella catarhalis infection; BASB034; diagnosis; staging;
KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW hearing loss; antibacterial drug.
XX
OS Moraxella catarhalis.
XX
PN WO200015802-A1.
XX
PD 23-MAR-2000.
XX
XX 14-SEP-1999; 99WO-EP06781.
PF
XX 14-SEP-1998; 98GB-0020002.
PR
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
PA
XX Ruelle J;
PI
XX WPI: 2000-271440/23.
DR N-PSDB; AAA10700.
XX
XX Novel BASB034 polynucleotides and polypeptides from Moraxella
PT catarhalis used to prepare vaccines against bacterial infections
PS Claim 3; Fig 2; 106pp: English.
XX
XX This sequence represents a Moraxella catarhalis BASB034 polypeptide from
CC strain Mc2931 (ATCC 43617). The invention relates to BASB034 polypeptides
CC from M. catarhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The
CC BASB034 polynucleotides and polypeptides may be employed as research
CC reagents and material for the discovery of treatments and diagnostics for
CC diseases, particularly human diseases. They are particularly used to
CC diagnose and treat M. catarhalis infections. They can be used for
CC diagnosis of disease, staging of disease, or determining response of an
CC infectious organism to drugs. The polynucleotides may be used as a source
CC for hybridization probes, and for screening of genetic mutations,
CC serotype, organism or strain identification, identification of mutations
CC in BASB034 sequences, and as components of arrays which are useful for
CC diagnostic and prognostic purposes. The polypeptides can be used to
CC produce antibodies. The polypeptides can also be used in vaccine
CC formulations, and to identify agonists and antagonists. The polypeptides,
CC antibodies, agonists and antagonists (which are bacteriostatic) are used
CC for the treatment and prevention of diseases such as otitis media in
CC infants and children, pneumonia in elderlies, sinusitis, nosocomial
CC infections and invasive diseases, and chronic otitis media with hearing
CC loss. The polypeptides, agonists and antagonists are also used for
CC screening of antibacterial drugs. The BASB034 products of the invention
CC can be used screen for new antibacterial compounds that may target
CC resistant bacteria.
XX
XX Sequence 442 AA:
SO
Query Match 50.7%; Score 224; DB 21; Length 442;
Best Local Similarity 100.0%; Pred. No. 6.6e-209;
Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 219 QFTPNFEFRAPELKFQVSVKAKAEDLWGTDSJLWFGYTQOSHMOJFNKNSRPFVHYQ 278
DB 219 qftpnfeafapelkfqsvvkaaedlwtgdsdlwtfgytqshwqjfnknsrpfvhyq 278

QY 279 PEIPLTQPVYSDLPWDGKVRNIGMGAVHHSNGESAKLSRSNRAVYLMAGMEKNLTVMPR 338
Db 279 PEIPLTQPVYSDLPWDGKVRNIGMGAVHHSNGESAKLSRSNRAVYLMAGMEKNLTVMPR 338
QY 339 IWGRIFKEGSGSQPPDDNPDLIDYRGYGDVREPLYOLENKSNI SGTVRYNPRSGKALOLDY 398
Db 339 Iwgrlfkegsgsqppddnpdlldyygydvrflyqlenksnls gtlvrynprrsgkaldy 398
QY 399 VYPLGKISGYFQIFGCGSLIDYNHBA TSFGVGLMLNDMMGL 442
Db 399 VYPLGKISGYFQIFGCGSLIDYNHBA TSFGVGLMLNDMMGL 442

Search completed: May 3, 2002, 10:23:31
Job time: 252 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2002, 10:24:14 ; Search time 36.69 Seconds
(without alignments)
271.094 Million cell updates/sec

Title: US-09-787-083-6

Perfect score: 442
Sequence: 1 MKVSLSTLTLSLCPAILA.....YNHEATFGVGLMDWGL 442

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 212252 seqs, 22503292 residues

Word size : 15

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCUS.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

No matches found

Search completed: May 3, 2002, 10:24:14
Job time: 185 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 30, 2001, 14:26:28 ; Search time 77.9 Seconds
(without alignments)
829.941 Million cell updates/sec

Title: US-09-787-083-8
Perfect score: 2363
Sequence: 1 MKVSLSTLTSLPCFALIA.....YNHEATSGVGLMLNDWGL 442

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP:archaea:*
2: SP:bacteria:*
3: SP:fungi:*
4: SP:human:*
5: SP:invertebrate:*
6: SP:mammal:*
7: SP:mhc:*
8: SP:organelle:*
9: SP:phage:*
10: SP:plant:*
11: SP:rodent:*
12: SP:virus:*
13: SP:vertebrate:*
14: SP:unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	815	34.5	382	2	09K0U7 neisseria m
2	815	34.5	409	2	09J1T1 neisseria m
3	372	15.7	306	2	09CL22 pasteurilla
4	366	15.5	329	2	09PMU8 campylobact
5	360.5	15.3	292	2	09SID7 yersinia ps
6	360	15.2	289	2	09LE69 salmonella
7	356.5	15.1	292	2	09Z4N8 enterobacte
8	342.5	14.5	297	2	03Z349 campylobact
9	246.5	10.4	355	2	025241 helicobacte
10	240.5	10.2	355	2	09ZLX5 helicobacte
11	123	5.2	278	2	09XB53 erwilia car
12	117	5.0	1686	4	000443 homo sapien
13	111.5	4.7	602	11	063485 rattus norv
14	108.5	4.6	821	2	059241 bacillus sp
15	107.5	4.5	824	2	09F216 bacillus sp
16	106.5	4.5	901	5	018749 caenorhabd1
17	106	4.5	798	4	09UN32 homo sapien
18	106	4.5	798	4	09UBK2 homo sapien
19	103	4.4	435	2	09X9C0 streptococc

20	102.5	4.3	739	5	09V9E6 drosophila
21	102.5	4.3	1046	2	084941 streptococc
22	101.5	4.3	403	5	P91736 hydra magni
23	101.5	4.3	783	2	043554 bacillus sp
24	100.5	4.3	660	2	09K6K8 bacillus ha
25	100.5	4.3	1509	11	061194 mus musculu
26	100.5	4.3	1658	11	061182 mus musculu
27	100.5	4.3	3247	12	065553 mus musculu
28	99.5	4.2	797	2	09RDW6 bovine herp
29	99	4.2	422	2	09RCZ6 lactobacill
30	99	4.2	5005	2	09RCZ6 streptomyc
31	98.5	4.2	467	4	09P2P5 ureaplasma
32	98.5	4.2	467	4	09UJM5 homo sapien
33	98.5	4.2	788	10	09BRY2 homo sapien
34	98.5	4.2	838	10	09SCV5 arabidopsis
35	97.5	4.1	683	2	09AC815 arabidopsis
36	96.5	4.1	749	2	09ABU7 caulobacter
37	96	4.1	791	10	049137 anaerocellu
38	96	4.1	4307	5	019319 caenorhabd1
39	95.5	4.0	576	2	09L115 streptomyc
40	95.5	4.0	772	2	09Z3V2 pseudomonas
41	95.5	4.0	1372	10	09FL92 arabidopsis
42	95	4.0	523	4	09H856 homo sapien
43	95	4.0	765	2	054183 streptomyc
44	95	4.0	871	4	09Y5C3 homo sapien
45	95	4.0	938	4	09Y5F7 homo sapien

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	382 AA.
09K0U7	09K0U7			
AC	09K0U7			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	PHOSPHOLIPASE A1, PUTATIVE.			
GN	NMB0464.			
OS	Neisseria meningitidis (serogroup B).			
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.			
OX	NCBI_TaxID=491;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MC58 / SEROGROUP B;			
RX	MEDLINE=20175755; PubMed=10710307;			
RA	Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,			
RA	Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,			
RA	Nelson W.C., Gwinn M.L., Deboy R., Peterson J.D., Hickey E.K.,			
RA	Hait D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,			
RA	Mason T., Ciecko A., Parksey D.S., Blair E., Cifton H., Clark E.B.,			
RA	Collon M.D., Ullrichback T.R., Khouri H., Qin H., Vamathevan J.,			
RA	Gill J., Scarlato V., Maignan V., Pizzi M., Grandi G., Sun L.,			
RA	Smith H.O., Fraser C.M., Moxon E.R., Rappelli R., Venter J.C.,			
RT	"Complete genome sequence of Neisseria meningitidis serogroup B strain			
RL	MC58."			
RL	Science 287:1809-1815(2000).			
EMBL	AE002403; AAF40901.1; -			
DR	EMBL: NMB0464; -			
DR	TI:G: NMB0464; -			
DR	InterPro: IPR003187; PLAI.			
DR	Pfam: PF02253; PLAI; 1.			
KW	Complete proteome.			
SO	SEQUENCE 382 AA; 42714 MW; B468A802F062E836 CRC64;			

Query Match 34.5%; Score 815; DB 2; Length 382;
Best Local Similarity 44.7%; Pred. No. 8.7e-63;
Matches 163; Conservative 67; Mismatches 111; Indels 24; Gaps 8;
Qy 87 INCSALNODIMRLACYDTLVHGETPAVI-----KTKSRIRLDETIVQFI-KGRPOVYQOE 140
Db 33 LQCALALDNTVRLACVYRIFRAQUPSSAGGEGESKAVLNLTEIVRSRLDKGEAVIYVEK 92

Qy	141	TTTPIFLMGNGKMLTKTKDAQQLYAAKQFPLSLSDLDLRNN-TPLMSSRPHPNPVLP	199
Db	93	GGDAI-----PADSAGETADYIPUSLMTDLDDNDRGLGVRHHNPMLMP	1399
Qy	200	IFMHGKPNRSNPSPSH-EAKQFTPNERRAPELKQVSVKVAADMLGTSDDIMEFGYTOQ	258
Db	140	LMYNNSPNYAGSPSPTRGTYTOQEKFGQOQRAETKQLQVSEFKSIADDEKTRADLMFGYTOR	199
Qy	259	SHMOLEFN-GKNSRPFRRVHDYQPELFTLQPVYSDLPWQCKKRMVGMGVHNSNGESAKLSR	317
Db	200	SDMOQLYNNGRSAPFRRNDYQPELFTLQPVYKADLPFGGRMLMLACGVHOSNOQSPRESR	259
Qy	318	SMNRAYTLMAAGEMKNKLTFMPRIPIRKEFGSGSOPDDNPDLIDYGGDYVRFYTOLEKNS	377
Db	260	SMNRITVMAAGEMKGLTYIPVWVRAPDQ-SGDK-NDNPDLADYMGVGDYKLYQRLNDRQ	317
Qy	378	NISGTVRYNPPRSKGALQLDVYVPLGKIGISGYFOIFQGYGOSLIDYVHEATSFQVGLMLN	437
Db	318	NVYSVLRNPNPTKGALTEAATYTFPIKGLKGVVGFHGYGSESLIDYNNHKGNGIGIGLMFN	377
Qy	438	DWNGI 442	
Db	378	DLDGI 382	
RESULT	2		
Q9J721			
09J721	PRELIMINARY:	PRT:	409 AA.
AC	09J721:		
DT	01-OCT-2000 (TREMblrel, 15, Created)		
DT	01-OCT-2000 (TREMblrel, 15, Last sequence update)		
DT	01-JUN-2001 (TREMblrel, 17, Last annotation update)		
DE	PURATIVE PHOBOLIPASE.		
CN	NMA2021.		
OS	Neisseria meningitidis (serogroup A).		
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.		
OX	NCBI_taxid=65699;		
RN	[1]		
RC	SEQUENCE FROM N.A.		
RP	STRAIN=22491 / SEROGROUP A / SEROTYPE 4A;		
RX	MEDLINE=20222556; PubMed=10761919;		
RA	Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,		
RA	Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,		
RA	Dayes R.M., Davis P., Devlin K., Feltham T., Hamlin N., Holtroyd S.,		
RA	Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,		
RA	Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,		
RA	Whitehead S., Spratt B.G., Barrall B.G.; a strain of Neisseria		
RT	"Complete DNA sequence of a serogroup A strain of Neisseria		
RT	meningitidis 22491."		
RL	Nature 404:502-506(2000).		
DR	EMBL: AL162757; CAB85240.1; .		
DR	InterPro: IPR003187; Pfam1.		
DR	Pfam: PF02253; Pfam1.		
KM	Complete proteome.		
SO	SEQUENCE 409 AA; 45862 MW; CD6565B064D01A41 CRC64;		

	Query Match	34.5%	Score 815;	DB 2;	Length 409;
	Best Local Similarity	44.7%;	Pred.	No. 9,6e-63;	
	Matches 163;	Conservative	67;	Mismatches 111;	Indels 24; Gaps 8
Oy	87	INCSALNODIMKLACYDYLHGETPAVI-----KTKRSIRLDETIIQWTI-KGKPQVAYOE	140		
		: : :: : : : : : : : : : :			
Dd	60	LQCAALTDVTFLLACDRIFAAOLPSSAQEGEESAVALNTFTTVASSLIDKGAVIAYEK	119		
Oy	141	TTPDFILMGNEKGLTKKDAKOLEVAAKOFTPLSLFSDLDRRN-TPLWSSRPHPMYVLP	199		
		: : : : : : : : : : : : : : : :			
Dd	120	GGDAL-----PADSAGETADITYTPLSMYDLDDKNDLRGLLGREHNPMYTMP	166		
Oy	200	IEMHCKPNRSPMPSH-EAKOFTPNFRAPRLKFQVSYSVKAABEDLMGDSDLMEGTQQ	258		
		:: :: : : : : : : : : : : : : : : : :			
Dd	147	LWNNSPNAPASPPRGTTVOEKGQOKAEATKLYQSFSEKIAAIEDLFKTRADLMFGTYTOR	226		

OY	259	SHMOJFN-GKSRSPFRVNDXQPELEFLTORPVSDLPMPGKVMIEMGAVHNSNGESATLSR	3147
		: : : : : : : : : : : : : : : : : : : : : : : : :	
Db	227	SDMQITINGGRKSAPFRNTDYPRELFILQPVKADLPFGGRRLMLAGVYHNSNGOSRPESR	286
OY	318	SMNRATYIAGMGEMKRLVTMPRIWGRIFKEGSGSQPDNDPILDDYVYGDVRFYQLDENKS	377
		: : : : : : : : : : : : : : : : : : : : : : : : :	
Db	287	SMNRATYIAGMGEMKRLVTIPVWVRAFDQ-SGDK-NQNPDLADYMGCDVKYQRLNDRQ	344
OY	378	NISGTVARNPSPSGALOLDVYVYLKKGISGYPDIPGCGOSLIDYVHHEATSPGVGLMN	437
		: : : : : : : : : : : : : : : : : : : : : : : : :	
Db	345	NVYSVLRNPXTGIGALTAATFFPKKGLKGVNGRFGHGEGLIDYVHKNKGISGILMFN	404
OY	438	DMMG 442	
		: : :	
Db	405	DLDGI 409	
RESULT	3		
O9CL22			
ID	O9CL22	PRELIMINARY;	PRT: 306 AA.
AC	O9CL22:		
DT	01-JUN-2001 (TREMblrel_17, Created)		
DT	01-JUN-2001 (TREMblrel_17, Last sequence update)		
DT	01-JUN-2001 (TREMblrel_17, Last annotation update)		
DE	HYPOTHEITICAL PROTEIN PM1426.		
GN	PM1426.		
OS	Pasteurella multocida.		
OC	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;		
CC	Pasteurella.		
OX	NCBI_TaxId=747;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=PM70;		
RX	MEDLINE=21145866; PubMed=11248100;		
RA	May B.J., Zhang Q., Li L.L., Paustian M.L., Whitlam T.S., Kapur V.;		
RT	"Complete genomic sequence of Pasteurella multocida PM70."		
RL	Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).		
DR	EMBL: AE006179; AAK03510.1; -		
DR	InterPro: IPR003187; PLAI1.		
DR	Pfam: PF02253; PLAI1.1.		
SW	Hypothetical protein; Complete proteome.		
SEQUENCE	306 AA; 35580 MW; EAF3DE8C1C22B26E CRC64;		

[illegible]

DE PHOSPHOLIPASE A (EC 3.1.1.32).
 GN PLDA OR C71351.
 OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
 OC Campylobacter.
 NC NCB1_TaxID=197;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCCTC 11168;
 RX MEDLINE=20150912; PubMed=10688204;
 RA Parkhill J., Wren B.W., Mungall K., Kellley J.M., Churcher C.,
 RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,
 RA Jagers K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
 RA Quail M.A., Rajandream M.A., Rutherford K.M., Van Vliet A.H.M.,
 RA Whitehead S., Barrett B.G.;
 RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
 RT reveals hypervariable sequences.";
 RL Nature 403:665-668(2000).
 RL EMBL: AL139078; CAB73778.1;
 DR InterPro: IPR003187; Pfam: 1.
 DR Pfam: PF02253; PLAI: 1.
 DR PRINTS: PR01486; PHPLIPASEA1.
 KW Complete proteome.
 SQ SEQUENCE 329 AA; 38880 MW; A32731F2B751AC44 CRC64;

Query Match 15.5%; Score 366; DB 2; Length 329;
 Best Local Similarity 33.2%; Pred. No. 9.2e-24;
 Matches 105; Conservative 48; Mismatches 125; Indels 38; Gaps 13;

OY 130 IKGPKQVYVQETDPIFLMGKNEKMLTKKAKOLEYAAKQTPISLSPDLDNRNT--PLW 187
 DB 43 LKSSSVLISOBONSSQATOTONSTITKEKO-----DSRLALNLYGENSEPNPLG 96
 OY 188 SSRPHNMYVLPFMHCKPNRSPPTPSHAKQFTPNFRAPELKFQVSVKKAEDLMT 247
 DB 97 IS-SYKNMYFLP-FAYSFNSLGVANNKSEA-----KFOLSYKRLLENLGL 141
 OY 248 DSDLMFYTOOSHWOITNGKNSRPFVHDYOPETFLTOPYV-SLPLMDGKVRMIGMAVH 306
 DB 142 DEKYIYITQTSWQIY--EHSSPFRFTNYOPEFIDLPLTKDYEFENMLR--VGILH 196
 OY 307 HSNGESAK--LSRSMNAYILMAGMEKMLNLYVPRMGRIFRFGSSGSDPDNDLIDYGY 364
 DB 197 ESNCKGCEMTQSRSMNRIYVSTALILNKFVPRMRYR---PNNKDDDPALILHMGN 253
 OY 365 GDVRFYOLEKNSNISGTVRYNPR--SGKALQIDYVYPL-GKISGFQYFQGGOSLI 421
 DB 254 FDVNLAV-LDDDFYINLMLNNKLFHNNKGAIOVDLGYDIFNNGIYWILOYFNGYGESLI 312
 OY 422 DYNHEATSFVGLMLN 437
 DB 313 DYNKHLQRLSTGLIS 328

RESULT 5
 OQSID7 PRELIMINARY; PRT: 292 AA.
 AC OQSID7;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE PHOSPHOLIPASE A.
 GN PLDA.
 OS Yersinia pseudotuberculosis.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Yersinia.
 NC NCB1_TaxID=633;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YPIII PIB1.
 RA Karlyshev A.V., Minzeler E.A., Williams K.J., Oyston P.C.,
 RA Tibball R.W., Wren B.W.;

RT "Biochip-based Signature-Tagged Mutagenesis: identification and
 RT characterisation of Y. pseudotuberculosis gene plda essential for
 RT virulence in mice.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ245393; CAB51586.1;
 DR InterPro: IPR003187; PLAI: 1.
 DR Pfam: PF02253; PLAI: 1.
 DR PRINTS: PR01486; PHPLIPASEA1.
 SQ SEQUENCE 292 AA; 33758 MW; BE712D908ACB6BA5 CRC64;

Query Match 15.3%; Score 360.5; DB 2; Length 292;
 Best Local Similarity 39.2%; Pred. No. 2.3e-23;
 Matches 83; Conservative 34; Mismatches 86; Indels 9; Gaps 5;

OY 229 ELKPOVSKVKAEDLMTGSDLMFGYTOOSHWOIFNGKNSRPFVHDYOPETFLTOPV- 287
 DB 86 EVKFOLSLAFPIWGRIGADNLSLGSYTORSMQANSSESSPFRFTNYEOPFLAMSTD 145
 OY 288 YSDLPMDGKVRMIGMAVHHNSGESAKLSRSMNAYILMAGMEKMLNLYVPRMGRIFRFG 347
 DB 146 YELAGW--TFREVEFGFNHOSNGKADPTSRSMNRYTRVMAQRGNLEIDLKPMYRIEESD 203
 OY 348 GSGQPDNDLIDYGYGDVRFYOL--ENKSNISGTVRYNPRSGKALQIDYVPLKGI 406
 DB 204 S---KDNFPIITKMGYRLKLVGIALGDSVSLDG--RYNNMTGCGAEMKSTPIIRHV 258
 OY 407 SGYFOIFQYGOQLIDYVHEATSFVGLMLND 438
 DB 259 RFTYQVFSYGESMIDYFNFRQTRVGVGIMLND 290

RESULT 6
 OQ16N9 PRELIMINARY; PRT: 289 AA.
 AC OQ16N9;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE PLDA PROTEIN.
 GN PLDA.
 OS Salmonella typhimurium LT2.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 NC NCB1_TaxID=99287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SGSC1412;
 RA Washu;
 RT "The Salmonella typhimurium Genome Sequencing Project.";
 RT submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SGSC1412;
 RA Waterston R.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF233324; AAF3435.1;
 DR InterPro: IPR003187; PLAI: 1.
 DR Pfam: PF02253; PLAI: 1.
 SQ SEQUENCE 289 AA; 32967 MW; DA97F5E1651C49C6 CRC64;

Query Match 15.2%; Score 360; DB 2; Length 289;
 Best Local Similarity 39.7%; Pred. No. 2.6e-23;
 Matches 85; Conservative 29; Mismatches 90; Indels 10; Gaps 4;

OY 226 RAEPLKPOVSKVKAEDLMTGSDLMFGYTOOSHWOIFNGKNSRPFVHDYOPETFLTO 285
 DB 83 RKDEYKFOLSLAFPIWGRIGILGPNISVLSASTOKSMQOLSNKSESPFRFTNYEOPFLIGF 142
 OY 286 PV-YSDLPMDGKVRMIGMAVHHNSGESAKLSRSMNAYILMAGMEKMLNLYVPRMGRIF 344
 DB 143 ATDYRFAGW--TLRDYEVKGYHDSNGRSDPTSRSMNRLYTRLMADENGMLVLEVAPMYVI- 199

RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Mallin E.,
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
 RA Venter J.C.;
 RT "The complete genome sequence of the gastric pathogen *Helicobacter*
 RT *pylori*.";
 RL Nature 388:539-547(1997).
 DR EMBL: AE000564; AAD07564.1; -
 DR TIGR: HP0499; -
 DR InterPro: IPR003187; PLA1.
 DR Pfam: PF02253; PLA1; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 355 AA; 42486 MW; 461960F15E65AB0A CRC64;

Query Match 10.4%; Score 246.5; DB 2; Length 355;
 Best Local Similarity 25.1%; Pred. No. 2,7e-13;
 Matches 82; Conservative 41; Mismatches 117; Indels 87; Gaps 11;

QY 157 KDAKOLEYAKOFTPLSLSPDLDNRNTPLMSSRPHNPMVLPFEMGKPNRSPNTPSHE 216
 DB 69 KKLNMMDYLGTYFLPRYHSF-----TFIFQVYHPINP----- 102
 QY 217 AKOFTNEFAPRLKFOVSVYKAAEDLMGTSDLMFGYTQOSHMOJFNKNSRPFVHD 276
 DB 103 ---YORNEF-----KFOISFRVPEFRHILMTKGLYLATYQTMFOJLYNDPOSAIPRMN 154
 QY 277 YQEIFLTPQVYSDLPWDGKY---RMIGCAVHNSG-ESAKLSRSNNRAYLMAGMEKN 332
 DB 155 FMPFLIYVYPI-NKPFQGGKIGNFESEIWMQHSNGVGACQYQPFNK---EGNPENQ 209
 QY 333 LTVMPRI-----WRIKFGSGSQP-----DNPDP 357
 DB 210 FPGQPIYKDYNGQKDYRMGCRSVSAGQRPVFLWEKGLKIMVAYWPYDQSNPN 269
 QY 358 ILDYGEGDYRFLY-----OLENKSINISGTVRYNPRSGKALQLDYVYPLKGISGYE 410
 DB 270 LIDYMGYGNMKIDYRGRHHFELQYDIFQYWRD--RWHGAFRLGYTRINPFV 327
 QY 411 QIFQIGYQSILIDYNHEATSFVGLMLN 437
 DB 328 QWFNGYGDGLYEYDFVFSNRIGVGIRLN 354

RESULT 10
 Q9ZLX5 PRELIMINARY; PRT; 355 AA.
 AC Q9ZLX5;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE PUTATIVE PHOSPHOLIPASE A1.
 GN PLDA OR JHP0451.
 OS *Helicobacter pylori* J99 (Campylobacter *pylori* J99).
 OC Bacteria; Proteobacteria; epsilon subdivision; *Helicobacter* group;
 OC *Helicobacter*.
 OX NCBI_Taxid=85963;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99120557; PubMed=9923682;
 RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
 RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human
 RT gastric pathogen *Helicobacter pylori*.";
 RL Nature 397:176-180(1999).
 DR EMBL: AE001479; AAD06029.1; -
 DR InterPro: IPR003187; PLA1.
 KW Complete proteome.
 SQ SEQUENCE 355 AA; 42329 MW; B3CE9810EBA2FDC9 CRC64;

Query Match 10.2%; Score 240.5; DB 2; Length 355;
 Best Local Similarity 25.1%; Pred. No. 8.9e-13;
 Matches 83; Conservative 44; Mismatches 109; Indels 95; Gaps 13;

QY 157 KDAKOLEYAKOFTPLSLSPDLDNRNTPLMSSRPHNPMVLPFEMGKPNRSPNTPSHE 216
 DB 69 KKLNMMDYLGTYFLPRYHSF-----TFIFQVYHPINP----- 102
 QY 217 AKOFTNEFAPRLKFOVSVYKAAEDLMGTSDLMFGYTQOSHMOJFNKNSRPFVHD 276
 DB 103 ---YORNEF-----KFOISFRVPEFRHILMTKGLYLATYQTMFOJLYNDPOSAIPRMN 154
 QY 277 YQEIFLTPQVYSDLPWDGKY---RMIGCAVHNSG-ESAKLSRSNNR----- 321
 DB 155 FMPFLIYVYPI-NKPFQGGKIGNFESEIWMQHSNGVGACQYQPFNKEGNPENQPPQ 213
 QY 322 -----AYLMAG-----MEWK--NLTVMPRIKGRIFKFGSGSQPDP- 354
 DB 214 PVIYKDYNGQKDYRMGCRSVSAGNALCFVLWEKGLKIMVAYWPY-----PYDQ 265
 QY 355 -NPDILYGYGVYRFLY-----OLENKSINISGTVRYNPRSGKALQLDYVYPLKG 406
 DB 266 SNPQLDYMGYGNMKIDYRGRHHFELQYDIFQYWRD--RWHGAFRLGYTRINPFV 323
 QY 407 SGYFOIGYQSILIDYNHEATSFVGLMLN 437
 DB 324 GIYQWFGYGDGLYEYDFVFSNRIGVGIRLN 354

RESULT 11
 Q9XB53 PRELIMINARY; PRT; 278 AA.
 AC Q9XB53;
 DT 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
 DE KDU1.
 GN *Erwinia carotovora*.
 OS *Erwinia carotovora*.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Pectobacterium.
 OX NCBI_Taxid=554;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 39048; GS101;
 RX MEDLINE=98065591; PubMed=9402024;
 RA McGowan S.J., Sebahia M., O'Leary S., Hardie K.R., Williams P.,
 RA Stewart G.S., Bycroft B.W., Salmon G.P.;
 RT "Analysis of the carboxenem gene cluster of *Erwinia carotovora*:
 RT definition of the antibiotic biosynthetic genes and evidence for a
 RT novel beta-lactam resistance mechanism.";
 RL Mol. Microbiol. 26:545-556(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 39048; GS101;
 RX MEDLINE=98276484; PubMed=9614345;
 RA McGowan S.J., Bycroft B.W., Salmon G.P.;
 RT "Bacterial production of carboxenems and clavams: evolution of beta-
 RT lactam antibiotic pathways.";
 RL Trends Microbiol. 6:203-208(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 39048; GS101;
 RA McGowan S.J.;
 RT Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U17224; AAD38237.1; -
 SQ SEQUENCE 278 AA; 31341 MW; 63769F4F3550E1B7 CRC64;

Query Match 5.2%; Score 123; DB 2; Length 278;
 Best Local Similarity 21.6%; Pred. No. 0.011;
 Matches 74; Conservative 52; Mismatches 128; Indels 88; Gaps 16;


```

QY 23 QAAVNPVAFVDEVSSENDLGODNELPIDVQSATOSTDTANPLDEHEPELYTTALEN 82
DB 81 ELERLNQVAVLTERKKELETLTAODRNGI-----OSQFTRAKEELEAEKRLRT---N 131
QY 83 KTLALINCSALNODIMRLACVDTLVHGETPAVVIKTRISIRLDETIWQ--TTKGRQVYVOE 140
DB 132 ERLSQEVEYLTEDEVKRL--NEKLAKESNT---TKGELQKLDELQASVTVYKREKRLBOE 186
QY 141 -----TTDPFLMGNEKG-----MLTKDAKOLEYAAKQFTPLS 174
DB 187 KELLHQNQSMNLNTELTCTDELLALGREGNEILELCTLENKEEDAISSHSSASPSA 246
QY 175 LSTDLDNRNTPFLMSSRRHNPMYVLPFMHCKPKNRSPPTPSHAKQFTPNRPAELKFOY 234
DB 247 LS--SSPNLSPTQMSQKTP-----VPAQREKAPSGSQEKKKIRPRGORDSSYYWEI 298
QY 235 SVKVKAAEDLMTGDSLMFGYTOOSH-----QIFNGKNSRPRVHDYOPEI----- 281
DB 299 EASEWMLSTRIGSGS---FSTVYKGGKHHGDVAVKILKVDPDTPQLOAFRNEVAVLKRTR 355
QY 282 -----FLTQ-----PYSDL--PMDGKVRMI-----GMGAVNH 307
DB 356 HVNILLFMGYMTKDNLAIVQWCGSSLYKHLHYQETKQFQWFLIDIAQTAQMDYLHA 415
QY 308 SNGESAKLSKSNRATVYLAAGMEK-----NLTVMPRIKGRIFKEGSGSO----- 351
DB 416 KNIIHRDKMS--NNIFLHEBLTYKIGDFGLATYKSRM-----SGSQOVEDPTGSYLMM 466
QY 352 -----PDGNPDIL--DYGYGDVRFLYOLENKSNIQYRYNPRSGKALQDLYYPL 402
DB 467 APEYTRKQDNNPFSFQSDVSYGLIV--LYEL-----MTGELPYSHINNROI-----IFMV 515
QY 403 KKG 405
DB 516 GRG 518

```

```

DB 9 LISSILLIVLLSL-----PPTALABGNTREDNEKHLGNDVYKRPSEAGALQLOEYWD 63
QY 65 ANPL-DEHEPELY-----TTALENKTMLINCSAL-----NODIMRLACV-DTLVHGET 110
DB 64 QMTLVDOHGEKIQOLRGSMTHGLQWPELINDNAYKALANDMESNMIRLAMYGENGVASN 123
QY 111 PAVIKTRISIRLDETI-----WQF-----IKGPOVY 137
DB 124 PELIKSRVIKIGDIALIENDMTVIYDMVHAPGDRPDPRYACADEFFRDIAALYRNPPIIT 183
QY 138 YQETTDV-----IFLNGNEKGM/TKKDAKOLEYAAKQFTPLSFDLDRN---NTP 186
DB 184 YELANEPSSNNNGGAGIPNNEGNNAVK-----EADPTVEMLRDSGNADNIIIVGSPN 238
QY 187 WSSRP-----HNDVYLPITF--MHGKPNRS--PNTPSHAKQFTPNRPAELKFT 232
DB 239 WSORPDLAADNPIDDHMTYVHYETGSHAATESYPRPFTNSERGNVMSYTRVA---L 294
QY 233 QVSVKVKAAEDLMCT-----DSDLMPGYTOO-----SHMOLFNGKNS-----R 270
DB 295 ENGVAVEFATE--WGTQOANGDGGEYFDEADVWIEFLNENNISWAMWSLITN--KNEVSGAFT 351
QY 271 PFVYH-----DYQPELITQPYSDLPW--- 293
DB 352 PFELGKSMTSLDGPDPQVWVPEELSLSGEYVYRARIKGVNER--IDRTKYTKVLMDFN 408
QY 294 DGKVMIGMGAVHSHNGES-----AKLSRSMNRAVYLAAGMEKNTLYMPPIW 340
DB 409 DGTQKGFV-----NGDSPVEDVYIENEGALKLSIGDASNDVSEGNVANNARLSADGW 462
QY 341 GR 342
DB 463 GK 464

```

```

RESULT 14
ID 059241 PRELIMINARY; PRT; 821 AA.
AC 059241:
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE ENDO-1,4-BETA-GLUCANASE (EC 3.2.1.4) (CELLULOSE) (ENDOGLUCANASE)
DE (CARBOXYMETHYL CELLULOSE).
OS Bacillus sp.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1409;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KSM-64;
RA Sumitomo N., Ozaki K., Ito S.;
RL Submitted (MAR-1993) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ENDODIOLYSIS OF 1,4-BETA-D-GLUCOSIDIC
CC LINKAGES IN CELLULOSE.
CC EMBL: M84963; AAA73189.1; -.
DR HSSP: 085465; 1A3H.
DR InterPro: IPR001547; Glyco_hydro_F5.
DR Pfam: PF00150; cellulase; 1.
DR PROSITE: PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN_1.
KW Hydrolase; Glycosidase.
SQ SEQUENCE 821 AA; 90910 MW; 73DA38FEF0B40B5C CRC64;

```

Query Match 4.5%; Score 107.5; DB 2; Length 824;
 Best Local Similarity 19.7%; Pred. No. 1.2;
 Matches 91; Conservative 56; Mismatches 147; Indels 167; Gaps 27;

Query Match 4.6%; Score 108.5; DB 2; Length 821;
 Best Local Similarity 19.9%; Pred. No. 0.94;
 Matches 96; Conservative 54; Mismatches 159; Indels 173; Gaps 25;

Query Match 4.5%; Score 107.5; DB 2; Length 824;
 Best Local Similarity 19.7%; Pred. No. 1.2;
 Matches 91; Conservative 56; Mismatches 147; Indels 167; Gaps 27;

```

Db 85 GLQWPEPEILNDNAYKALSNMDSNMIRLAMY---VGENGYATNPBLIKQRYIDGIELAI 140
QY 121 RLDETI---WQT-----IKGKPOVYQOETTD-----IF 146
Db 141 ENDWYVIVDMVHVHAGDPDRDPRVYAGAKDPFREIALYFNNPHIITELANEPSSNNNGAG 200
QY 147 LMGNEKGLTKRDAKQLEYAAKQFTPLSLSPDLDRN---NTPLMSSRP-----H 192
Db 201 IPNNEEGWKAVK---EYADPIYEMLRKSGNADONIIIVGSPNWSORPDLAADNPIDDH 255
QY 193 NPMYVLPYLFMHGKPNRSPNTPSHKAKQFTPNFEPAPBLK-----FQVSYKKAADLMGT 247
Db 256 HTMYTVHFYTGSHAASTESPSE---TPNSEGNVMSNTRYALENGVAVPATE--WGT 308
QY 248 -----DSDLWFGYTQO-----SHMOIFNGKNS-----RPFYV----- 274
Db 309 SQASGDDGPTFDEADVWLEFLNENNISMANNSLTN--KNEVSGATFPFELGKSNATNLDPG 367
QY 275 --HDYQP-EIFLT-----QPV---YSDLPW--DGKVRMIGMGAVHHSNGE 311
Db 368 PDHVVAPPEELISGEYVRARIKGVNYPEIDRTKYTKVLMDFNDDGTQK--GFGVNSDSPNK 425
QY 312 SAKLSRSNBRAYLMAGME-----WKNLTVMPRIMGR 342
Db 426 ELIAVDNENNTLKVSGLDVSNVSDGDNFEMNARLSANGWK 466

```

Search completed: November 30, 2001, 14:26:30
 Job time: 567 sec

RESULT	1
AY85268	
ID	AA85268 standard; Protein; 442 AA.
XX	
AC	AA85268;
XX	
DT	29-JUN-2000 (first entry)
XX	
DE	BASB034 amino acid sequence #1.
XX	
KW	Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KW	vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW	sinusitis; nosocomial infection; invasive disease; chronic otitis media;
XX	hearing loss; antibacterial drug.
OS	Moraxella catarrhalis.
XX	
PN	WO200015802-A1.
XX	
PD	23-MAR-2000.
XX	
PF	14-SEP-1999; 99WO-EP06781.
XX	
PR	14-SEP-1998; 98GB-0020002.
XX	
PA	(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX	
PI	Ruelle J;
XX	
DR	WPI; 2000-271440/23.
DR	N-PSDB; AAA10700.
XX	
PT	Novel BASB034 polynucleotides and polypeptides from Moraxella
PT	catarrhalis used to prepare vaccines against bacterial infections

XX Claim 3; Fig 2; 106pp; English.
PS
CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC strain Mc2931 (ATCC 43617). The invention relates to BASB034 polypeptides
CC from M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2965. The
CC BASB034 polynucleotides and polypeptides may be employed as research
CC reagents and material for the discovery of treatments and diagnostics for
CC diseases, particularly human diseases. They are particularly used to
CC diagnose and treat M. catarrhalis infections. They can be used for
CC diagnosis of disease, staging of disease, or determining response of an
CC infectious organism to drugs. The polynucleotides may be used as a source
CC for hybridization probes, and for screening of genetic mutations.
CC serotype, organism or strain identification, identification of mutations
CC in BASB034 sequences, and as components of arrays which are useful for
CC diagnostic and prognostic purposes. The polypeptides can be used to
CC produce antibodies. The polypeptides can also be used in vaccine
CC formulations, and to identify agonists and antagonists. The polypeptides,
CC antibodies, agonists and antagonists (which are bacteriostatic) are used
CC for the treatment and prevention of diseases such as otitis media in
CC infants and children, pneumonia in elderlies, sinusitis, nosocomial
CC infections and invasive diseases, and chronic otitis media with hearing
CC loss. The polypeptides, agonists and antagonists are also used for
CC screening of antibacterial drugs. The BASB034 products of the invention
CC can be used screen for new antibacterial compounds that may target
CC resistant bacteria.
XX
SQ Sequence 442 AA;

Query Match 100.0%; Score 442; DB 21; Length 442;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKSSTLTLSITSCAIIAIOQAQAVNPVAFVDEVRSENDIGQDNELPIDVQSATQSA 60
DB 1 mkvsstltlsitscfaillaigqagvnpvafvdevrsendigqdnelpidvqsatqsa 60
QY 61 STDTANPLDEHEPELYTTALENKTMKLNCSALNODIMRLACYDTLVHGETPAVYIKTKRST 120
DB 61 stdtanpldehepelyttalenktmklncsalnqdimrlacydtlvhgetpaviktkrst 120
QY 121 RLDETTWQIKGKPOYIYVETTDPIFLMGNEKMLTKKAQOLEYAAKOTPTLSFDD 180
DB 121 rldettwqikgkpoiyvettddpiflmgnkmltkkaqoleyaakotptlsfdd 180
QY 121 rldettwqikgkpoiyvettddpiflmgnkmltkkaqoleyaakotptlsfdd 180
QY 181 RNNTPLMSSRPNNPMTVLPDIFMHGKPNRSPNTPSHBAKQFTPNFEPAPELKEFVSAYKVA 240
DB 181 rnntpmlssrpnnpmtvlpdifmhgkpnrsnpnpsheakqftpnfepapelkftvsaykva 240
QY 241 AEDLMCTSDLMFGYQOQSHWOJFNGKNSRPFVNDYDPEIFLTQPVYSDLPMDGKVRMT 300
DB 241 aedlmctsdlmfgyqoqshwojfnknsrpfvndydpelfltqpvysdelpmdgkvrm 300
QY 241 aedlmctsdlmfgyqoqshwojfnknsrpfvndydpelfltqpvysdelpmdgkvrm 300
QY 301 GMGAVHNSGESKLSRSNNRAYLTMAGMEKMLTVMPRIKGRIFKESGSGSDPDNDIID 360
DB 301 gmgavhnsgeksklsrsnnraylmgmewkmltvmprirkgrifkesgsgsdpdndiid 360
QY 361 YVGYGVRFLYOLENKSNTSGTVRNPYRSGKALQLDYVYPLGKSGFQIIOQYQGS 420
DB 361 yvgygvrfllyolenksntsgtvrnpysgkalgldyvyplgksgfqiioqyqgs 420
QY 421 IDYNHEATSEFGVGLMLNDMMGL 442
DB 421 idynheatsfgyglmlndmmgl 442

RESULT 2
AAV85271
ID AAV85271 standard; Protein; 442 AA.
XX
AC AAV85271;
XX

DT 29-JUN-2000 (first entry)
XX
DE BASB034 amino acid sequence #4.
XX
KW Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
XX hearing loss; antibacterial drug.
OS
XX Moraxella catarrhalis.
XX
PN WO200015802-A1.
XX
XX 23-MAR-2000.
XX
PD 14-SEP-1999; 99WO-EP06781.
PF
XX 14-SEP-1998; 98GB-0020002.
PR
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PA Ruelle J;
XX
PI
XX
DR WPI: 2000-271440/23.
DR N-PSDB; AAA10703.
XX
XX Novel BASB034 polynucleotides and polypeptides from Moraxella
XX catarrhalis used to prepare vaccines against bacterial infections -
PS Claim 3; Page 69; 106pp; English.

QY 1 MKSSTLTLSITSCAIIAIOQAQAVNPVAFVDEVRSENDIGQDNELPIDVQSATQSA 60
DB 1 mkvsstltlsitscfaillaigqagvnpvafvdevrsendigqdnelpidvqsatqsa 60
QY 61 STDTANPLDEHEPELYTTALENKTMKLNCSALNODIMRLACYDTLVHGETPAVYIKTKRST 120
DB 61 stdtanpldehepelyttalenktmklncsalnqdimrlacydtlvhgetpaviktkrst 120
QY 121 RLDETTWQIKGKPOYIYVETTDPIFLMGNEKMLTKKAQOLEYAAKOTPTLSFDD 180
DB 121 rldettwqikgkpoiyvettddpiflmgnkmltkkaqoleyaakotptlsfdd 180
QY 181 RNNTPLMSSRPNNPMTVLPDIFMHGKPNRSPNTPSHBAKQFTPNFEPAPELKEFVSAYKVA 240
DB 181 rnntpmlssrpnnpmtvlpdifmhgkpnrsnpnpsheakqftpnfepapelkftvsaykva 240
QY 241 AEDLMCTSDLMFGYQOQSHWOJFNGKNSRPFVNDYDPEIFLTQPVYSDLPMDGKVRMT 300
DB 241 aedlmctsdlmfgyqoqshwojfnknsrpfvndydpelfltqpvysdelpmdgkvrm 300
QY 301 GMGAVHNSGESKLSRSNNRAYLTMAGMEKMLTVMPRIKGRIFKESGSGSDPDNDIID 360
DB 301 gmgavhnsgeksklsrsnnraylmgmewkmltvmprirkgrifkesgsgsdpdndiid 360
QY 361 YVGYGVRFLYOLENKSNTSGTVRNPYRSGKALQLDYVYPLGKSGFQIIOQYQGS 420
DB 361 yvgygvrfllyolenksntsgtvrnpysgkalgldyvyplgksgfqiioqyqgs 420
QY 421 IDYNHEATSEFGVGLMLNDMMGL 442
DB 421 idynheatsfgyglmlndmmgl 442

Query Match 74.0%; Score 327; DB 21; Length 442;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 427; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 15 CFAIIAIOQAQAVNPVAFVDEVRSENDIGQDNELPIDVQSATQSA 74
DB 15 cfaiillaioqaavnpvafvdevrsendigqdnelpidvqsatqsa 74
QY 75 LYTTALENKTMLINCALNODIMRLACYDTLVHGETPAVYIKTKRSTRLDETTWQIKGK 134
DB 75 lyttalenktmlincsalnqdimrlacydtlvhgetpaviktkrst rldettwqikgk 134
QY 135 QVITYOETTDPIFLMGNEKMLTKKAQOLEYAAKOTPTLSFDDLRNNTPLMSSRPNNP 194
DB 135 qvityoettdpiflmgnkmltkkaqoleyaakotptlsfddlrnntplmssrpnnp 194

QY 195 MYVLPIFMHGKPNRSPNTPSHAKOFTPNFRAPRLKFOVSKYKAAEDIMGTDSDIAPG 254
 DB 195 MYVLPIFMHGKPNRSPNTPSHAKOFTPNFRAPRLKFOVSKYKAAEDIMGTDSDIAPG 254
 QY 255 YTOQSHMOIFNGKNSRPRVDYOPETLTOFPVYSDLPWDGKVRMIGMGAHNSGSEAK 314
 DB 255 YTOQSHMOIFNGKNSRPRVDYOPETLTOFPVYSDLPWDGKVRMIGMGAHNSGSEAK 314
 QY 315 LSRSMNRAVLMAGMEKMLTYVPRIMGRIFKEGSGSQPDNDIIDYGYGDFRFLYOLE 374
 DB 315 LSRSMNRAVLMAGMEKMLTYVPRIMGRIFKEGSGSQPDNDIIDYGYGDFRFLYOLE 374
 QY 375 NKSNTSGVRNPRSGKALOLDIYPRIGKISGTFQIFQCGGSLIDYNHEATSPGVL 434
 DB 375 NKSNTSGVRNPRSGKALOLDIYPRIGKISGTFQIFQCGGSLIDYNHEATSPGVL 434
 QY 435 MLNDMGL 442
 DB 435 MLNDMGL 442
 RESULT 3
 ID AAY85269 standard; Protein: 442 AA.
 AC AAY85269;
 DT 29-JUN-2000 (first entry)
 DE BASB034 amino acid sequence #2.
 XX Moraxella catarrhalis infection; BASB034: diagnosis; staging;
 KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
 KM sinusitis; nosocomial infection; invasive disease; chronic otitis media;
 KW hearing loss; antibacterial drug.
 XX Moraxella catarrhalis.
 OS WO200015802-A1.
 PN 23-MAR-2000.
 PD 23-MAR-2000.
 PE 14-SEP-1999; 99WO-EP06781.
 PR 14-SEP-1998; 98GB-0020002.
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 PI Ruelle J;
 XX WPI: 2000-271440/23.
 DR N-PSDB: AAA10701.
 PT Novel BASB034 polynucleotides and polypeptides from Moraxella
 catarrhalis used to prepare vaccines against bacterial infections -
 Claim 3; Page 67; 106pp; English.
 This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 strain Mc2908. The invention relates to BASB034 polypeptides from
 M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
 polynucleotides and polypeptides may be employed as research reagents and
 material for the discovery of treatments and diagnostics for diseases,
 particularly human diseases. They are particularly used to diagnose and
 treat M. catarrhalis infections. They can be used for diagnosis of
 disease, staging of disease, or determining response of an infectious
 organism to drugs. The polynucleotides may be used as a source for
 hybridization probes, and for screening of genetic mutations, serotype,
 organism or strain identification, identification of mutations in BASB034
 sequences, and as components of arrays which are useful for diagnostic
 and prognostic purposes. The polypeptides can be used to produce
 antibodies. The polypeptides can also be used in vaccine formulations,

CC and to identify agonists and antagonists. The polypeptides, antibodies,
 CC agonists and antagonists (which are bacteriostatic) are used for the
 CC treatment and prevention of diseases such as otitis media in infants and
 CC children, pneumonia in elders, sinusitis, nosocomial infections and
 CC invasive diseases, and chronic otitis media with hearing loss. The
 CC polypeptides, agonists and antagonists are also used for screening of
 CC antibacterial drugs. The BASB034 products of the invention can be used
 CC screen for new antibacterial compounds that may target resistant
 CC bacteria.
 XX Sequence 442 AA;
 SQ
 Query Match 50.7%; Score 224; DB 21; Length 442;
 Best Local Similarity 100.0%; Pred. No. 9.2e-209;
 Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 219 QFTPNFRAPRLKFOVSKYKAAEDIMGTDSDIAPGTYTOQSHMOIFNGKNSRPRVDYQ 278
 DB 219 QFTPNFRAPRLKFOVSKYKAAEDIMGTDSDIAPGTYTOQSHMOIFNGKNSRPRVDYQ 278
 QY 279 PEIFLTPQVYSDLPWDGKVRMIGMGAHNSGSEAKLSRSMNRAVLMAGMEKMLTYVPR 338
 DB 279 PEIFLTPQVYSDLPWDGKVRMIGMGAHNSGSEAKLSRSMNRAVLMAGMEKMLTYVPR 338
 QY 339 IMGRIFKEGSGSQPDNDIIDYGYGDFRFLYOLENKSNTSGTYRVNPRSGKALOLDY 398
 DB 339 IMGRIFKEGSGSQPDNDIIDYGYGDFRFLYOLENKSNTSGTYRVNPRSGKALOLDY 398
 QY 399 VYPLGKISGYFQIFQCGGSLIDYNHEATSPGVLMLNDMGL 442
 DB 399 VYPLGKISGYFQIFQCGGSLIDYNHEATSPGVLMLNDMGL 442
 RESULT 4
 ID AAY85270 standard; Protein: 442 AA.
 AC AAY85270;
 DT 29-JUN-2000 (first entry)
 DE BASB034 amino acid sequence #3.
 XX Moraxella catarrhalis infection; BASB034: diagnosis; staging;
 KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
 KM sinusitis; nosocomial infection; invasive disease; chronic otitis media;
 KW hearing loss; antibacterial drug.
 XX Moraxella catarrhalis.
 OS WO200015802-A1.
 PN 23-MAR-2000.
 PD 23-MAR-2000.
 PE 14-SEP-1999; 99WO-EP06781.
 PR 14-SEP-1998; 98GB-0020002.
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 PI Ruelle J;
 XX WPI: 2000-271440/23.
 DR N-PSDB: AAA10702.
 PT Novel BASB034 polynucleotides and polypeptides from Moraxella
 catarrhalis used to prepare vaccines against bacterial infections -
 Claim 3; Page 68; 106pp; English.
 This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
 strain Mc2913. The invention relates to BASB034 polypeptides from

PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX
 DR WPI; 2000-062150/05.
 DR N-PSDB; AA253919.
 XX
 PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics -
 XX
 PS Claim 2; Page 904; 1453pp; English.
 XX
 CC AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941
 CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA255473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 XX
 SQ Sequence 370 AA;

Query Match 2.0%; Score 9; DB 21; Length 370;
 Best Local Similarity 100.0%; Pred. No. 2,3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 PEFLTPGV 287
 |||||||||
 Db 209 pelfltqpv 217

RESULT 7
 AAY75158
 ID AAY75158 standard; Protein; 370 AA.
 XX
 AC AAY75158;
 XX
 DT 21-MAR-2000 (first entry)
 XX
 DE Neisseria meningitidis ORF 582 protein sequence SEQ ID NO:1790.
 XX
 KM Neisseria meningitidis: Neisseria gonorrhoeae; antigen; vaccine;
 KM antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
 KM antibacterial; gene therapy.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO9957280-A2.
 XX
 PD 11-NOV-1999.
 XX
 PF 30-APR-1999; 99WO-US09346.
 XX
 PR 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.

PR 25-FEB-1999; 99US-0121528.
 XX (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX
 DR WPI; 2000-062150/05.
 DR N-PSDB; AA253920.
 XX
 PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics -
 XX
 PS Claim 2; Page 905; 1453pp; English.
 XX
 CC AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941
 CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA255473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 XX
 SQ Sequence 370 AA;

Query Match 2.0%; Score 9; DB 21; Length 370;
 Best Local Similarity 100.0%; Pred. No. 2,3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 PEFLTPGV 287
 |||||||||
 Db 209 pelfltqpv 217

RESULT 8
 AAY70629
 ID AAY70629 standard; Protein; 374 AA.
 XX
 AC AAY70629;
 XX
 DT 18-JUL-2000 (first entry)
 XX
 DE Neisseria meningitidis serogroup B strain H44/76 BASB033 protein.
 XX
 KM BASB033; diagnosis; prophylaxis; treatment; antibacterial; vaccine;
 KM Neisseria meningitidis infection.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO200015801-A1.
 XX
 PD 23-MAR-2000.
 XX
 PF 09-SEP-1999; 99WO-EP06718.
 XX
 PR 14-SEP-1998; 98GB-0020003.
 XX
 PA (SMIR) SMITHKLINE BEECHAM BIOLOGICALS.
 XX
 PI Ruelle J;
 XX
 DR WPI; 2000-271439/23.
 DR N-PSDB; AA252134.

PT Isolated BASB033 polypeptides and polynucleotides of *Neisseria*
PT meningitidis, useful for diagnosis, prophylaxis and treatment of *N.*
PT meningitidis infection -
XX
XX
PS Claim 4: Page 59; 93pp; English.
XX
CC The present sequence is a BASB033 protein from
CC *Neisseria meningitidis* serogroup B strain H44/76. The protein
CC shows homology to the *Klebsiella pneumoniae* outer membrane
CC phospholipase A. The present sequence is useful for diagnosis,
CC prophylaxis and treatment of *N. meningitidis* infection. It may also be
CC used for the discovery and development of antibacterial compounds and
CC in vaccine compositions.
XX
SQ Sequence 374 AA;

Query Match 2.0%; Score 9; DB 21; Length 374;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 PEIFLRQPV 287
Db 213 pefltpv 221

RESULT 9
AAV70628
ID AAV70628 standard; Protein; 375 AA.
XX
XX AAV70628;
XX
DT 18-JUL-2000 (first entry)
XX
DE *Neisseria meningitidis* serogroup B strain ATCC13090 BASB033 protein.
XX
XX BASB033; diagnosis, prophylaxis; treatment; antibacterial; vaccine;
XX *Neisseria meningitidis* infection.
XX
OS *Neisseria meningitidis*.
XX
PN WO200015801-A1.
XX
PD 23-MAR-2000.
XX
PF 09-SEP-1999; 99WO-EP06718.
XX
PR 14-SEP-1998; 98GB-0020003.
XX
PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Ruelle J;
XX
DR MPI: 2000-271439/23.
DR N-PSDB; AA52133.
XX
PT Isolated BASB033 polypeptides and polynucleotides of *Neisseria*
PT meningitidis, useful for diagnosis, prophylaxis and treatment of *N.*
PT meningitidis infection -
XX
PS Claim 4: Page 58; 93pp; English.
XX
XX The present sequence is a BASB033 protein from
XX *Neisseria meningitidis* serogroup B strain ATCC13090. The protein
XX shows homology to the *Klebsiella pneumoniae* outer membrane
XX phospholipase A. The present sequence is useful for diagnosis,
XX prophylaxis and treatment of *N. meningitidis* infection. It may also be
XX used for the discovery and development of antibacterial compounds and
XX in vaccine compositions.
SQ Sequence 375 AA;

Query Match 2.0%; Score 9; DB 21; Length 375;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 PEIFLRQPV 287
Db 214 pefltpv 222

RESULT 10
AAG11137
ID AAG11137 standard; Protein; 278 AA.
XX
XX AAG11137;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 9741.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
PN EP103405-A2.
XX
PD 06-SEP-2000.
XX
XX
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140685.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 25-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-SEP-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 1.88; Score 8; DB 21; Length 278;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Gaps 0;
 Qy 174 slsfdldr 181
 |||||
 Db 271 slsfdldr 278

RESULT 11
ACG1136
ID ACG1136 standard; Protein; 283 AA.
XX ACG1136;
AC ACG1136;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 9740.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148655.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.

PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 8; DB 21; Length 283;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 SLSEFDLR 181
| | | | | | | | | |
Db 276 SLSEFDLR 283

RESULT 12
AA01135
ID AA01135 standard; Protein; 299 AA.
XX
AC AA01135;

XX 17-OCT-2000 (first entry)
DT Arabidopsis thaliana protein fragment SEQ ID NO: 9739.
XX
DE
XX Protein Identification: signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0135494.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139494.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.

```

PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.

```

```

PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 1.8%; Score 8; DB 21; Length 299;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 174 SLSEFLDR 181
Db 292 slsfldr 299

```

```

RESULT 13
AAR52127
ID AAR52127 standard; Peptide: 16 AA.
XX
AC AAR52127;
XX
XX 27-SEP-1996 (first entry)
DT
XX
XX Mouse light chain surface patch S03410.
DE
XX
XX antibody; humanised; murine; human; heavy chain; light; variable;
KW framework region; complementarity determining region; reshaping;

```


KW modelling; surface residue; modify.
XX
OS Mus sp.
XX
PN EP592106-A1.
XX
PD 13-Apr-1994.
XX
PF 07-SEP-1993; 93EP-0307051.
XX
PR 09-SEP-1992; 92US-0942245.
XX
PA (PEDE/) PEDERSEN J T.
PA (IMMU-) IMMUNOGEN INC.
XX
PI Guild BC, Pedersen JT, Rees AR, Roguska MA, Searle SMJ;
XX
DR WPI; 1994-120230/15.
XX
PT Method of resurfacing of rodent antibodies to produce humanised
PT antibody forms - for producing non-human antibodies with improved
PT therapeutic efficiency by presenting human surface on V-region
XX
PS Example 1; Page 14; 230pp; English.
XX
CC Modification of a rodent antibody or fragment by resurfacing in order
CC to produce a humanised rodent antibody can be determined by calculating
CC homology between murine and human antibody surfaces. In order to test
CC the resurfacing approach of the invention, three humanisation
CC experiments were set up. (1) traditional loop grafting; (2) resurfacing
CC approach using most similar chain; and (3) resurfacing approach using
CC human sequences with most similar surface residues. AAR52069-159 are the
CC surface residue patterns in mouse light chain antibody variable regions.
CC These "patches" were used in the third method, where rodent light and
CC heavy chains were matched and the most similar human sequence found
CC independently only over the surface residues indicated in AAR52030-67.
XX
SQ Sequence 16 AA:

Query Match 1.6%; Score 7; DB 15; Length 16;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 387 PRSGKA 393
Db 4 prsgkga 10
|||||

RESULT 14
AAG76069
ID AAG76069 standard; Protein; 33 AA.
XX
AC AAG76069;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen protein SEQ ID NO:6833.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma.
XX
OS Homo sapiens.
XX
PN WO20012920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26524.
XX
PR 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX

PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR WPI; 2001-235357/24.
XX
DR N-PDB; AAH35474.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
PS Claim 11; Page 8285-8286; 9803pp; English.
XX
CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ Sequence 33 AA:

Query Match 1.6%; Score 7; DB 22; Length 33;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 LTLSTLS 14
Db 22 ltlstls 28
|||||

RESULT 15
AAY79483
ID AAY79483 standard; peptide; 50 AA.
XX
AC AAY79483;
XX
DT 01-AUG-2000 (first entry)
XX
DE Rat alpha tubulin variant peptide (aa314-363).
XX
KW Rat; alpha tubulin; mutant; mutein; variant;
KW metabolic X syndrome; hypertension; stroke; diabetes;
KW insulin resistance; obesity; dyslipidemia; anorectic;
KW hypotensive; cerebroprotective; vasodilator; antidiabetic;
KW antilipemic; gene therapy; spontaneously hypertensive; SHR.
XX
OS Rattus sp.
XX
FH Key
FH Peptide
FT 1..50
FT /note- "corresponds to residues 314-363 of
FT full-length protein"
FT Misc-difference 27
FT /note- "replaces Thr in wild-type"
XX
PN WO200018918-A2.
XX
PD 06-APR-2000.
XX

PF 28-SEP-1999; 99WO-US22494.
 XX
 PR 28-SEP-1998; 98US-0161939.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA;
 XX
 DR WPI; 2000-303450/26.
 XX
 PT Novel genes encoding e.g. CD36, SGLT2, and kynurenine aminotransferase,
 PT which are predictive and therapeutic for stroke, hypertension, diabetes
 PT and obesity
 XX
 PS Claim 2; Fig 4G; 79pp; English.
 XX
 CC This sequence represents amino acids 314-363 of alpha-tubulin
 CC from a spontaneously hypertensive rat (SHR). This region of
 CC alpha-tubulin carries a Ser for Thr amino acid substitution when
 CC compared to the corresponding region in control Wistar Kyoto (WKY)
 CC rats (see AAY/9484). The alpha-tubulin gene is differentially
 CC expressed in SHR, SHR-stroke prone (SP) and WKY rats. It may be
 CC involved in a predisposition to vascular injury. The SHR rat
 CC is an animal model for human metabolic X syndrome. Genes encoding
 CC sodium dependent glucose cotransporter, kynurenine aminotransferase,
 CC CD36, aldolase A, atrial natriuretic factor, alpha-cardiac myosin
 CC and alpha-tubulin were identified as being potentially associated
 CC with hypertension, obesity and insulin resistance. These
 CC proteins, including isolated proteins comprising the present
 CC sequence, are used for treating, preventing and diagnosing
 CC ischemic and metabolic diseases and disorders, such as stroke,
 CC hypertension, diabetes and obesity, especially insulin resistivity,
 CC dyslipidemia and ischemic stroke (all claimed). Polynucleotides
 CC encoding them may also be used in gene therapy and antisense
 CC therapy protocols.
 XX
 SQ Sequence 50 AA;

Query Match 1.6%; Score 7; DB 21; Length 50;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 114 IKTKRSI 120
 |||||
 DB 22 Iktkrsi 28

Search completed: May 9, 2002, 12:54:58
 Job time: 513 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:56:32 ; Search time 44.37 Seconds
(without alignments)
758.827 Million cell updates/sec

Title: US-09-787-083-4

Perfect score: 442
Sequence: 1 MKVSLSTLTLSTLPCFAILA.....YVHEATSPGVGLMDMGL 442

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	2.0	382	2 E81195	phospholipase A1,
2	9	2.0	409	2 H81831	probable phospholip
3	8	1.8	286	2 B36971	outer membrane pho
4	8	1.8	299	2 T47989	RAV-like protein -
5	8	1.8	1396	2 S36851	L-shaped tail fibre
6	7	1.6	122	2 E75377	hypothetical prote
7	7	1.6	155	2 A29659	beta-lactoglobulin
8	7	1.6	158	2 E86498	pts IIA protein (I
9	7	1.6	158	2 E72124	tubulin alpha-1 ch
10	7	1.6	197	2 A54506	hypothetical prote
11	7	1.6	219	2 T19438	tubulin alpha chain
12	7	1.6	240	2 A61544	coat protein - pot
13	7	1.6	264	2 A44959	coat protein - pot
14	7	1.6	267	2 A60366	genome polypotein
15	7	1.6	267	2 S14001	capsid protein - p
16	7	1.6	267	2 S26630	coat protein - pot
17	7	1.6	267	2 JCL527	tail fiber protein
18	7	1.6	267	2 S13239	abc transport ATP-
19	7	1.6	270	2 S73734	transaminase B hom
20	7	1.6	273	2 T36237	probable phosphor
21	7	1.6	288	2 E72576	probable ABC trans
22	7	1.6	292	2 T27985	hypothetical prote
23	7	1.6	306	2 T27985	probable RAV-like
24	7	1.6	310	2 T02280	polypotein - pota
25	7	1.6	313	2 JT0960	hypothetical prote
26	7	1.6	324	2 T23876	protease inhibit
27	7	1.6	326	2 A43939	hypothetical prote
28	7	1.6	338	2 F69437	hypothetical prote
29	7	1.6	342	2 G64411	hypothetical prote

30	7	1.6	342	2 G64394	hypothetical prote
31	7	1.6	349	2 T43920	ytic protein [impo
32	7	1.6	357	2 A45619	ubiquinol-cytochr
33	7	1.6	370	2 G83219	probable dihydroli
34	7	1.6	379	2 S13556	genome polypotein
35	7	1.6	381	2 G82132	ribonuclease D WC1
36	7	1.6	382	2 S13237	tail fiber protein
37	7	1.6	384	2 S70638	tubulin alpha chain
38	7	1.6	384	2 S70641	tubulin alpha chain
39	7	1.6	384	2 S70640	tubulin alpha chain
40	7	1.6	387	2 S33517	tubulin alpha chain
41	7	1.6	407	2 T19895	hypothetical prote
42	7	1.6	415	2 H70477	UDP-N-acetylmutamo
43	7	1.6	419	2 F85064	hypothetical prote
44	7	1.6	419	2 T25377	hypothetical prote
45	7	1.6	423	1 UBFA	tubulin alpha-1 ch

ALIGNMENTS

RESULT 1

E81195 phospholipase A1, probable NMB0464 [imported] - Neisseria meningitidis (strain MC58 s

C:Species: Neisseria meningitidis

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C:Accession: E81195

R:Telletlin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.

ri, H.; Qiu, H.; Yamathavan, J.; Gill, J.; Scariato, V.; Maignani, V.; Piazza, M.

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappunli, R.;

A>Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755

A:Accession: E81195

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-382 <TEP>

A:Cross-references: GB:AE002403; GB:AE002098; NID:g7225688; PID:NMB040901.1; PID:g7222

A:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: NMB0464

Query Match 2.0%, Score 9, DB 2; Length 382;

Best Local Similarity 100.0%; Pred. No. 0.51;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 279 PEIPLTPV.287

Db 221 PEIPLTPV.229

RESULT 2

H81831 probable phospholipase NMA2021 [imported] - Neisseria meningitidis (strain Z2491 serog

C:Species: Neisseria meningitidis

C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C:Accession: H81831

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo

l; Holtroyd, S.; Jagers, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre

Nature 404, 502-506, 2000

A>Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491

A:Reference number: A81775; MUID:20222556

A:Accession: H81831

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-409 <PAR>

A:Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PID:NCAB85240.1; PID:g738

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: NMA2021

Query Match 2.0%; Score 9; DB 2; Length 409;
 Best Local Similarity 100.0%; Pred. No. 0.54;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEFLTPPV 287
 |||||
 Db 248 PEFLTPPV 256

RESULT 3
 B36971

outer membrane phospholipase A (EC 3.1.1.-) precursor - *Klebsiella pneumoniae*
 C:Species: *Klebsiella pneumoniae*
 C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 18-Jun-1999
 C:Accession: B36971; S40129
 R:Brok, R.G.P.M.; Brinkman, E.; van Boxtel, R.; Bekkers, A.C.A.P.A.; Verheij, H.M.; Tomm
 J. Bacteriol. 176, 861-870, 1994
 A:Title: Molecular characterization of enterobacterial *pldA* genes encoding outer membran
 A:Reference number: A36971; MUID:94131966
 A:Accession: B36971
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-286 <BRO>
 A:Cross-references: EMBL:X76901; NID:g436880; PIDN:CAA54223.1; PID:g436881
 A:Note: authors translated the codon AAG for residue 112 as Arg
 C:Genetics:
 A:Gene: *pldA*
 C:Superfamily: bacterial phospholipase A1
 C:Keywords: carboxylic ester hydrolase

Query Match 1.8%; Score 8; DB 2; Length 286;
 Best Local Similarity 100.0%; Pred. No. 4.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 431 GVGILMND 438
 |||||
 Db 277 GVGILMND 284

RESULT 4
 T47989
 RAY-like protein - *Arabidopsis thaliana*
 N:Alternate names: protein F21F14.140
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T47989
 R:Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Catolico, L.; Artiguenave, F.; Se
 submitted to the Protein Sequence Database, February 2000
 A:Reference number: Z24481
 A:Accession: T47989
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-299 <CHO>
 A:Cross-references: EMBL:AL138642
 A:Experimental source: cultivar Columbia; BAC clone F21F14
 C:Genetics:
 A:Map position: 3
 A:Note: F21F14.140

Query Match 1.8%; Score 8; DB 2; Length 299;
 Best Local Similarity 100.0%; Pred. No. 4.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 SLSPDLR 181
 |||||
 Db 292 SLSPDLR 299

RESULT 5
 S36851

L-shaped tail fiber protein - phage T5
 N:Alternate names: ltf protein
 C:Species: phage T5
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 17-Nov-2000
 C:Accession: S65934; S01984; S36851
 R:Kallman, A.V.; Kulshin, V.E.; Shlyupnikov, M.G.; Ksenzenko, V.N.; Kryukov, V.M.
 FEBS Lett. 366, 46-48, 1995
 A:Title: The nucleotide sequence of the bacteriophage T5 ltf gene.
 A:Reference number: S65934; MUID:95309401
 A:Accession: S65934
 A:Molecule type: DNA
 A:Residues: 1-1396 <KAL>
 A:Cross-references: EMBL:X69460; NID:g15415; PIDN:CAA9220.1; PID:g15416
 R:Kallman, A.V.; Kryukov, V.M.; Bayev, A.A.
 Nucleic Acids Res. 16, 6230, 1988
 A:Title: The nucleotide sequence of bacteriophage T5 DNA at the region between early
 A:Reference number: S01982; MUID:88289370
 A:Accession: S01984
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 934-985, 'A', 987-1396 <KA2>
 A:Cross-references: EMBL:X07559
 C:Genetics:
 A:Gene: ltf
 C:Keywords: late protein; tail fiber

Query Match 1.8%; Score 8; DB 2; Length 1396;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 QSATQSAS 61
 |||||
 Db 87 QSATQSAS 94

RESULT 6
 E75377
 hypothetical protein - *Deinococcus radiodurans* (strain R1)
 C:Species: *Deinococcus radiodurans*
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: E75377
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
 M.; Shen, M.; Yamathayan, J.J.; Lam, P.; McDonald, L.; Uitterback, T.; Zalewski, C.;
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
 A:Reference number: A75250; MUID:20036896
 A:Accession: E75377
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-122 <WHI>
 A:Cross-references: GB:AE002002; GB:AE000513; NID:g6459345; PIDN:AAF1150.1; PID:g645
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR1580
 A:Map position: 1

Query Match 1.6%; Score 7; DB 2; Length 122;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 EHEPELY 76
 |||||
 Db 115 EHEPELY 121

RESULT 7
 A29699
 beta-lactoglobulin - eastern gray kangaroo
 C:Species: *Macropus giganteus* (eastern gray kangaroo)
 C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 18-Jul-1997

C:Accession: A29659
R:Godovac-Zimmermann, J.; Shaw, D.
Biol. Chem. Hoppe-Seyler 368, 879-886, 1987
A:Title: The primary structure, binding site and possible function of beta-lactoglobulin
A:Reference number: A29659; MUID:87299024
A:Accession: A29659
A:Molecule type: protein
A:Residues: 1-155 <GOD>
C:Superfamily: lipocalin; lipocalin homology
F:10-155/Domain: lipocalin homology <LIP>

Query Match 1.6%; Score 7; DB 2; Length 155;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 38 RSKNDLG 44
|||||||
DB 5 RSKNDLG 11

RESULT 8
E86498
Pts IIA Protein [Imported] - Chlamydomonas pneumoniae (strain J138)
C:Species: Chlamydomonas pneumoniae, Chlamydia pneumoniae
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: E86498
R:Shirai, M.; Hiraoka, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349
A:Accession: E86498
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-158 <STO>
A:Cross-references: GB:BA000008; NID:98978434; PIDN:BAA98271.1; GSPDB:GN00142
C:Experimental source: strain J138
C:Genetics:
A:Gene: ptnN_1
C:Superfamily: phosphotransferase system enzyme II; phosphotransferase system mannitol-8

Query Match 1.6%; Score 7; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 LSTLTLS 11
|||||||
DB 126 LSTLTLS 132

RESULT 9
E72124
pts iia protein - Chlamydomonas pneumoniae (strain CML029)
C:Species: Chlamydomonas pneumoniae, Chlamydia pneumoniae
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
C:Accession: E72124
R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606
A:Accession: E72124
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-158 <ARN>
A:Cross-references: GB:AE001591; GB:AE001363; NID:94376311; PIDN:AAD18213.1; PID:9437631
C:Experimental source: strain CML029
C:Genetics:
A:Gene: ptnN_1
C:Superfamily: phosphotransferase system enzyme II; phosphotransferase system mannitol-8
F:28-156/Domain: phosphotransferase system mannitol-specific enzyme II factor III homolog

Query Match 1.6%; Score 7; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 LSTLTLS 11
|||||||
DB 126 LSTLTLS 132

RESULT 10
A54506
tubulin alpha-1 chain - Plasmodium yoelii (fragment)
C:Species: Plasmodium yoelii
C>Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 13-Aug-1999
C:Accession: A54506
R:Arkella, R.; Arasu, P.; Vaidya, A.B.
Mol. Biochem. Parasitol. 30, 165-174, 1988
A:Title: Molecular clones of alpha-tubulin genes of Plasmodium yoelii reveal an unusu
A:Reference number: A54506; MUID:89014607
A:Accession: A54506
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-197 <AKE>
A:Cross-references: GB:M29816; NID:g160729; PIDN:AAA29779.1; PID:g160730
C:Superfamily: tubulin

Query Match 1.6%; Score 7; DB 2; Length 197;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTKRSI 120
|||||||
DB 82 IKTKRSI 88

RESULT 11
T19438
hypothetical protein C25A1.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T19438
R:Mortimore, B.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19124
A:Accession: T19438
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-219 <WIL>
A:Cross-references: EMBL:Z81038; PIDN:CAB02762.1; GSPDB:GN00019; CESP:C25A1.1
A:Experimental source: clone C25A1
C:Genetics:
A:Gene: CESP:C25A1.1
A:Map position: 1
A:introns: 33/1; 66/1; 151/2

Query Match 1.6%; Score 7; DB 2; Length 219;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 158 KDAKOLE 164
|||||||
DB 52 KDAKOLE 58

RESULT 12
A61544
tubulin alpha chain - common octopus (fragment)
C:Species: Octopus vulgaris (common octopus)
C>Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 10-Jul-1998
C:Accession: A61544
R:Zinov'eva, R.D.; Aleinikova, K.S.; Tomarev, S.I.

Dokl. Akad. Nauk SSSR 302, 462-467, 1988

A:Title: Isolation and structural characterization of cDNAs coding for alpha-tubulin of A:Reference number: A61544

A:Accession: A61544

A:Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-240 <2IN>

C:Superfamily: tubulin
F:239-240/Cleavage site: Glu-Tyr (tubulin-specific carboxypeptidase) #status predicted
F:239-240/Cross-Link: peptide (Glu-Tyr) (by tubulin-tyrosine ligase) #status predicted

Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 240;
Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTAKSI 120
|||
Db 124 IKTAKSI 130

RESULT 13

A44959

coat protein - potato virus Y (strain N) (fragment)

C:Species: potato virus Y, PVY

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Nov-2000

C:Accession: A44959

R:Hay, J.M.; Fellowes, A.P.; Timmerman, G.M.

Arch. Virol. 107, 111-122, 1989

A:Title: Nucleotide sequence of the coat protein gene of a necrotic strain of potato virus Y

A:Reference number: A44959; MID:90025790

A:Accession: A44959

A:Status: Preliminary

A:Molecule type: genomic RNA

A:Residues: 1-264 <HAY>

A:Cross-references: GB:M2470; NID:9333508; PIDN:AAA47182.1; PID:9333509

C:Superfamily: tobacco etch virus genome polypeptide

Query Match
Best Local Similarity 100.0%; Score 7; DB 2; Length 264;
Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 156 TKKDAKQ 162
|||
Db 8 TKKDAKQ 14

RESULT 14

A60366

coat protein - potato virus Y (strain TH)

C:Species: potato virus Y, PVY

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Nov-2000

C:Accession: A60366

R:Hataya, T.; Sano, T.; Ohshima, K.; Shikata, E.

Virus Genes 4, 339-350, 1990

A:Title: Polymerase chain reaction-mediated cloning and expression of the coat protein gene of potato virus Y

A:Reference number: A60366; MID:91143125

A:Accession: A60366

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-267 <HAT>

C:Superfamily: tobacco etch virus genome polypeptide

Query Match
Best Local Similarity 100.0%; Score 7; DB 2; Length 267;
Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 156 TKKDAKQ 162
|||
Db 11 TKKDAKQ 17

RESULT 15

S14001

genome polypeptide - potato virus Y (strain H) (fragment)

N:Contains: coat protein

C:Species: potato virus Y, PVY

A:Variety: strain H

C:Date: 21-Nov-1993 #sequence_revision 08-Nov-1996 #text_change 17-Nov-2000

C:Accession: S14001

R:Dalmay, T.; Balazs, E.

Nucleic Acids Res. 18, 6721, 1990

A:Title: Nucleotide sequence of an altered virulence potato virus Y coat protein gene

A:Reference number: S14001; MID:91067494

A:Accession: S14001

A:Status: translation not shown

A:Molecule type: genomic RNA

A:Residues: 1-267 <DAL>

A:Cross-references: EMBL:X54611; NID:961449; PIDN:CMA38432.1; PID:9330257

A:Experimental source: strain H

C:Superfamily: tobacco etch virus genome polypeptide

C:Keywords: coat protein; polypeptide

F:1-267/Product: coat protein #status predicted <CPR>

Query Match
Best Local Similarity 100.0%; Score 7; DB 2; Length 267;
Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 156 TKKDAKQ 162
|||
Db 11 TKKDAKQ 17

Search completed: May 9, 2002, 12:56:33
Job time: 603 sec

OY 431 GVGIMLND 438
 DB 277 GVGIMLND 284

RESULT 2

VLTF_BPRT5 STANDARD: PRT; 1396 AA.
 AC P1390: 048502;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE L-SHAPED TAIL FIBER PROTEIN (LTF PROTEIN).
 GN LTF.
 OS Bacteriophage T5.
 OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae.
 OX NCBI_TaxID=10726;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-95309401; PubMed-7789514;
 RA Kallman A.V., Kulshin V.E., Shlyapnikov M.G., Ksenzenko V.N.,
 RA Kryukov V.M.;
 RT "The nucleotide sequence of the bacteriophage T5 ltf gene.";
 RL FEBS Lett. 366:46-48(1995).
 RN (2)
 RP SEQUENCE FROM N.A.
 RA Kallman A.V.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN (3)
 RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
 RX MEDLINE-88289370; PubMed-3267228;
 RA Kallman A.V., Kryukov V.M., Bayev A.A.;
 RT "The nucleotide sequence of bacteriophage T5 DNA at the region
 RT between early and late genes.";
 RL Nucleic Acids Res. 16:6230-6230(1988).
 CC -1- FUNCTION: NONESSENTIAL PROTEIN THAT MEDIATES BINDING TO THE
 CC POLYMANNOSE O ANTIGEN.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL: X69460: CAA49220.1; -.
 DR EMBL: AJ001191: CAA04591.1; -.
 DR PIR: S01982; S01982.
 KW Late protein.
 FT CONFLICT 986 V -> A (IN REF. 2).
 SQ SEQUENCE 1396 AA; 147989 MW; 18CD2192P65FFC1 CRC64;

Query Match 1.8%; Score 8; DB 1; Length 1396;
 Best Local Similarity 100.0%; Pred. No. 6.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 54 QSATOSAS 61
 DB 87 QSATOSAS 94

RESULT 3

LACB_MACGI STANDARD: PRT; 155 AA.
 AC P11944;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE BETA-LACTOGLOBULIN.
 GN LGB.

OS Macropus giganteus (Eastern gray kangaroo).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
 OX NCBI_TaxID=9317;
 RN (1)
 RP SEQUENCE.

RX MEDLINE-87299024; PubMed-3620116;
 RA Godovac-Zimmermann J., Shaw D.;
 RT "Beta-lactoglobulin identified in marsupial milk. The primary
 RT structure, binding site and possible function of beta-lactoglobulin
 RT from eastern grey kangaroo (Macropus giganteus).";
 RL Biol. Chem. Hoppe-Seyler 368:879-886(1987).
 CC -1- FUNCTION: LACTOGLOBULIN IS THE PRIMARY COMPONENT OF WHEY, IT
 CC BINDS RETINOL AND IS PROBABLY INVOLVED IN THE TRANSPORT OF
 CC THAT MOLECULE.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SIMILARITY: BELONGS TO THE LIPOCALIN FAMILY.
 DR PIR: A29699; A29699.
 DR HSSP: P02754; 1BSQ.
 DR InterPro: IPR000566; Lipocln_cytfabp.
 DR Pfam: PF00061; Lipocalin; 1.
 DR PROSITE: PS00213; LIPOCALIN; FALSE NEG.
 KM Milk; whey; Retinol-binding; Transport; Lipocalin.
 FT DISULFID 61 153 BY SIMILARITY.
 FT DISULFID 103 115
 SQ SEQUENCE 155 AA; 18340 MW; 7BD7EA4A191530E5 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 155;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 38 RSKNDLG 44
 DB 5 RSKNDLG 11

RESULT 4

LACB_MACEU STANDARD: PRT; 174 AA.
 AC Q29614;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE BETA-LACTOGLOBULIN PRECURSOR.
 GN LGB OR BLG.
 OS Macropus eugenii (Tamar wallaby).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
 OX NCBI_TaxID=9315;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-91197370; PubMed-1707629;
 RA Collet C., Joseph R., Nicholas K.R.;
 RT "A marsupial beta-lactoglobulin gene: characterization and prolactin-
 RT dependent expression.";
 RL J. Mol. Endocrinol. 6:9-16(1991).
 RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-95314595; PubMed-7794241;
 RA Collet C., Joseph R.;
 RT "Exon organization and sequence of the genes encoding alpha-
 RT lactalbumin and beta-lactoglobulin from the tamar wallaby
 RT (Macropodidae, Marsupialia).";
 RL Biochem. Genet. 33:61-72(1995).
 RN (3)
 RP SEQUENCE OF 47-174 FROM N.A.
 RC TISSUE-Mammary gland;
 RX MEDLINE-95085593; PubMed-7993373;
 RA Collet C., Joseph R.;
 RT "The identification of nuclear and mitochondrial genes by sequencing
 RT randomly chosen clones from a marsupial mammary gland cDNA library.";
 RL Biochem. Genet. 32:181-190(1994).


```

CC -1- FUNCTION: LACTOGLOBULIN IS THE PRIMARY COMPONENT OF WHEY, IT
CC BINS RETINOL AND IS PROBABLY INVOLVED IN THE TRANSPORT OF
CC RETINOL MOLECULE.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LIPOCALIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L14959; AAA31597.1; -.
DR EMBL: L14954; AAA31597.1; JOINED.
DR EMBL: L14955; AAA31597.1; JOINED.
DR EMBL: L14956; AAA31597.1; JOINED.
DR EMBL: L14957; AAA31597.1; JOINED.
DR EMBL: L14958; AAA31597.1; JOINED.
DR EMBL: X15212; CAA33282.1; -.
DR HSSP: P02754; IBSO.
DR InterPro: IPR000566; Lipocln_cyFABP.
DR Pfam: PF00061; Lipocalin; 1.
DR PROSITE: PS00213; LIPOCALIN; FALSE NEG.
KW Milk; Whey; Retinol-binding; Transport; Lipocalin; Signal.
FT SIGNAL 1 18
FT CHAIN 1 18
FT DISULEID 19 174
FT DISULEID 79 172
FT DISULEID 122 134
FT CONFLICT 76 76
FT SEQUENCE 174 AA; 20219 MW; 6C5901051F99C991 CRC64;
SQ

```

```

Query Match 1.6%; Score 7; DB 1; Length 174;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 38 RSKNDLG 44
Db 23 RSKNDLG 29

RESULT 5
PRL_ICTPU STANDARD; PRT; 186 AA.
ID PRL_ICTPU
AC P51904;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE PROLACTIN (PRL).
OS Ictalarus punctatus (Channel catfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Siluriformes; Ictaluridae; Ictalurus.
OX NCBI_TaxID=7998;
RN [1]
RP SEQUENCE.
RC TISSUE=Pituitary; PubMed=1308206;
RX MEDLINE=93364578; Pubmed=1308206;
RA Watanabe K., Igarashi A., Noso T., Chen T.T., Dunham R.A.,
RA Kawachi H.;
RT "Chemical identification of catfish growth hormone and prolactin.";
RL Mol. Mar. Biol. Biotechnol. 1:239-249(1992).
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
DR HSSP: Q28632; IAN3.
DR InterPro: IPR001400; SOMATOTROPIN.
DR Pfam: PF00103; hormone; 1.
DR PRINTS: PR00836; SOMATOTROPIN.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW Hormone; Pituitary.

```

```

FT DISULEID 45 159
FT DISULEID 176 186
SQ SEQUENCE 186 AA; 20606 MW; 7BCCD89718F44E74 CRC64;

```

```

Query Match 1.6%; Score 7; DB 1; Length 186;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 41 NDLGQDN 47
Db 140 NDLGQDN 146

```

```

RESULT 6
TBA_PLAYO STANDARD; PRT; 197 AA.
ID TBA_PLAYO
AC P12543;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE TUBULIN ALPHA CHAIN (FRAGMENT).
OS Plasmodium berghei yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5862;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89014607; Pubmed=2459618;
RA Akella R., Arasu P., Valdivya A.B.;
RT "Molecular clones of alpha-tubulin genes of Plasmodium yoelii reveal
RT an unusual feature of the carboxy terminus.";
RL Mol. Biochem. Parasitol. 30:165-174(1988).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M29816; AAA29779.1; -.
DR PIR: A54506; A54506.
DR InterPro: IPR000217; Tubulin.
DR InterPro: IPR003008; Tubulin_FtsZ.
DR Pfam: PF00091; tubulin; 1.
DR PROSITE: PS00227; TUBULIN; PARTIAL.
KW Microtubules; GTP-binding.
FT NON_TER 1
FT SEQUENCE 197 AA; 22111 MW; 4DAAF199CC6CD319 CRC64;
SQ

```

```

Query Match 1.6%; Score 7; DB 1; Length 197;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 114 IKTKRSI 120
Db 82 IKTKRSI 88

```

```

RESULT 7
TBA_OCTVU STANDARD; PRT; 240 AA.
ID TBA_OCTVU
AC P24635;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)

```

DE TUBULIN ALPHA CHAIN (FRAGMENT).
 OS Octopus vulgaris (Octopus).
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;
 OC Incirrata; Octopodidae; Octopus.
 OX NCBI_TaxID=6645;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lens;
 RA Zinov'eva R.D., Aleinikova K.S., Tomarev S.I.;
 RT "Isolation and structural characterization of cDNAs coding for alpha-
 RT tubulin of the octopus eye lens."
 RL Dokl. Akad. Nauk SSSR 302:462-467(1988).
 CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
 CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
 CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
 CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
 CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to [license@isb-sid.ch](mailto:license@isb-sib.ch)).
 CC -----
 CC EMBL: X15845; CAA33844.1; -
 CC PIR: A61544; A61544.
 CC InterPro: IPR000217; Tubulin.
 CC InterPro: IPR003008; Tubulin_FtsZ.
 CC Pfam: PF00091; tubulin; 1.
 CC PROSITE: PS00227; TUBULIN; PARTIAL.
 CC Microtubules; GTP-binding.
 CC KW NON_TER
 CC FT SITE 240 240 INVOLVED IN POLYMERIZATION.
 CC SEQUENCE 240 AA; 26961 MW; 13BB3A1F740F2A16 CRC64;
 SO

Query Match 1.6%; Score 7; DB 1; Length 240;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTKRSI 120
 Db 124 IKTKRSI 130

RESULT 8
 ID Y304_MYCPN STANDARD; PRT; 270 AA.
 AC P75355;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN MG304 HOMOLOG
 DE (A05_ORF270L).
 GN MPN43 OR MP408.
 OS Mycoplasma pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 OC Mycoplasmales; Mycoplasma.
 OX NCBI_TaxID=2104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;
 RX MEDLINE=97105885; PubMed=8948633;
 RA Himmelfrich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
 RA Herrmann R.;
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
 RT pneumoniae."
 RT Nucleic Acids Res. 24:4420-4449(1996).
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
 CC (ABC TRANSPORTERS).
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to [license@isb-sid.ch](mailto:license@isb-sib.ch)).
 CC -----
 CC EMBL: AE000040; AAB96056.1; -
 CC InterPro: IPR003593; AAA.
 CC InterPro: IPR003439; ABC_transporter.
 CC InterPro: IPR001687; ATP_GTP_A.
 CC Pfam: PF00005; ABC_tran; 1.
 CC SMART: SM00382; AAA; 1.
 CC PROSITE: PS00211; ABC_TRANSPORTER; 1.
 CC KW Hypothetical protein; ATP-binding; Transport; Complete proteome.
 CC FT NP_BIND 36 43
 CC SEQUENCE 270 AA; 30770 MW; 6512640E4BC051B4 CRC64;
 SO

Query Match 1.6%; Score 7; DB 1; Length 270;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 168 KQFTPLS 174
 Db 125 KQFTPLS 131

RESULT 9
 ID LIVE_THEMA STANDARD; PRT; 273 AA.
 AC P74921;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PROBABLE BRANCHED-CHAIN AMINO ACID AMINOTRANSFERASE (EC 2.6.1.42)
 DE (BCAT).
 GN LIVE OR TM0831.
 OS Thermotoga maritima.
 OC Bacteria; Thermotogales; Thermotoga.
 OX NCBI_TaxID=2336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RX MEDLINE=99287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
 RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Utterback T.R., Malek J.A., Linner K.D., Garrett M.M.,
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 RT genome sequence of Thermotoga maritima."
 RT Nature 399:323-329(1999).
 RN [2]
 RP SEQUENCE OF 68-273 FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RX MEDLINE=97017137; PubMed=8863738;
 RA Guipaud O., Labedan B., Forterre P.;
 RT "A gyrB-like gene from the hyperthermophilic bacterium Thermotoga
 RT maritima."
 RT Gene 174:121-128(1996).
 CC -1- CATALYTIC ACTIVITY: L-LEUCINE + 2-OXOGLUTARATE = 4-METHYL-2-
 CC OXOPENTANOATE + L-GLUTAMATE (ALSO ACTS ON L-ISOLEUCINE AND
 CC L-VALINE).
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC -1- PATHWAY: VALINE AND ISOLEUCINE BIOSYNTHESIS.
 CC -1- SIMILARITY: BELONGS TO CLASS-IV OF PYRIDOXAL-PHOSPHATE-DEPENDENT
 CC AMINOTRANSFERASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AE001750; AAD35913.1; -
 DR EMBL: U96692; AAC44497.1; -
 DR HSSP: P00510; 1A3G.
 DR TIGR: TM0831; -
 DR InterPro: IPR001544; AminoTran_4.
 DR Pfam: PF01063; AminoTran_4; 1.
 DR ProDom: PD001961; AminoTran_4; 1.
 DR PROSITE: PS00770; AA-TRANSFER_CLASS_4; 1.
 DR Transferrase: AminoTransferase; Branched-chain amino acid biosynthesis;
 KW Pyridoxal phosphate; Complete proteome.
 FT CONFLICT 77 A -> R (IN REF. 2).
 FT SEQUENCE 273 AA; 3158 MW; 21628705612E90E3 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 273;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 172 PLSLSPD 178
 Db 61 PLSLSPD 67

RESULT 10
 MIAA_PSEPU STANDARD; PRT; 322 AA.
 AC 030762;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE TRNA DELTA(2) -ISOPENTENYL-PYROPHOSPHATE TRANSFERASE (EC 2.5.1.8) (TPP
 DE TRANSFERASE) (ISOPENTENYL-DIPHOSPHATE:TRNA ISOPENTENYLTRANSFERASE)
 DE (TPPase) (TPP).
 GN MIAA.
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M.
 RA Olekhnovich I.N., Gussin G.N.;
 RT "Attenuation of the Pseudomonas putida trpE and trpGDC genes.";
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CATALYZES THE FIRST STEP IN THE BIOSYNTHESIS OF
 CC 2-METHYLTHIO-N6-(DELTA(2)-ISOPENTENYL)-ADENOSINE (MS[2]116[a])
 CC ADJACENT TO THE ANTICODON OF SEVERAL TRNA SPECIES (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ISOPENTENYL DIPHOSPHATE + TRNA =
 CC PYROPHOSPHATE + TRNA CONTAINING 6-ISOPENTENYLADEOSINE.
 CC -1- SIMILARITY: BELONGS TO THE TPP TRANSFERASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF016312; AAB69443.1; -
 DR InterPro: IPR002627; TPPt.
 DR Pfam: PF01715; TPPt; 1.
 DR ProDom: PD004674; TPPt; 1.
 DR Transferrase: Nucleotidyltransferase; RNA processing; ATP-binding.
 KW NF_BIND 12 19 ATP (POTENTIAL).
 FT SEQUENCE 322 AA; 3548 MW; D1AA81A2D7B32F6A CRC64;

Query Match 1.6%; Score 7; DB 1; Length 322;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 FVDEVR 39
 Db 231 FVDEVR 237

RESULT 11
 SPL_BACBR STANDARD; PRT; 326 AA.
 AC P43131;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE PROTEASE INHIBITOR PRECURSOR (BBRPt).
 DE Bacillus brevis.
 OS Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Brevibacillus.
 OX NCBI_TaxID=1393;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-32; 104-112 AND 122-136.
 RC STRAIN=HPD31;
 RX MEDLINE=92304060; PubMed=1610177;
 RA Shiga Y., Hasegawa K., Tsuboi A., Yamagata H., Ueda S.;
 RT "Characterization of an extracellular protease inhibitor of Bacillus
 RT brevis HPD31 and nucleotide sequence of the corresponding gene.";
 RL Appl. Environ. Microbiol. 58:525-531(1992).
 CC -1- FUNCTION: SHOWS INHIBITORY ACTIVITY TOWARDS SERINE PROTEASES, SUCH
 CC AS TRYPSIN, CHYMOTRYPSIN, AND SUBTILISIN. MAY FORM A TRYPSIN-
 CC INHIBITOR COMPLEX IN A MOLAR RATIO OF 1:1. IT IS HEAT RESISTANT AT
 CC NEUTRAL AND ACIDIC PH.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- PTM: MAY UNDERGO ACTIVATION AFTER SECRETION. IT IS PRODUCED
 CC EXTRACELLULARLY IN MULTIPLE FORMS HAVING AT LEAST THREE DIFFERENT
 CC MOLECULAR WEIGHTS (BBRPt-A, -B, AND -C).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D10696; BAA01538.1; -
 KW Serine protease inhibitor; Protease inhibitor; Signal; Repeat.
 DR STGNAL 1 24
 FT CHAIN 25 326 SERINE PROTEASE INHIBITOR.
 FT CHAIN 104 326 SERINE PROTEASE INHIBITOR-C.
 FT CHAIN 104 7 SERINE PROTEASE INHIBITOR-B (POTENTIAL).
 FT CHAIN 122 326 SERINE PROTEASE INHIBITOR-A.
 FT DOMAIN 177 304 CONTAINS TWO APPROXIMATE REPEATS.
 FT REPEAT 177 208 1.
 FT REPEAT 272 304 2.
 SO SEQUENCE 326 AA; 35100 MW; 1C0456ABFA912E77 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 326;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 30 PVAFVDE 36
 Db 119 PVAFVDE 125

RESULT 12
 FLAD_METUA STANDARD; PRT; 342 AA.
 ID FLAD_METUA
 AC Q58305;


```

FT TRANSMEM 123 143 POTENTIAL.
FT TRANSMEM 150 170 POTENTIAL.
FT TRANSMEM 183 203 POTENTIAL.
FT TRANSMEM 217 237 POTENTIAL.
FT TRANSMEM 246 266 POTENTIAL.
FT TRANSMEM 278 298 POTENTIAL.
FT TRANSMEM 301 321 POTENTIAL.
FT TRANSMEM 331 351 POTENTIAL.
FT TRANSMEM 374 394 POTENTIAL.
SQ SEQUENCE 423 AA; 44322 MM; 166203BBB2A8A15B CRC64;

```

```

QY 114 IKTKRSI 120
Db 329 IKTKRSI 335

```

Search completed: May 9, 2002, 13:05:52
Job time: 607 sec

Query Match 1.6%; Score 7; DB 1; Length 423;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 402 LKKGISG 408
Db 104 LKKGISG 110

```

```

RESULT 15
TBA_ONCKE
ID TBA_ONCKE STANDARD; PRT; 444 AA.
AC P30436;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE TUBULIN ALPHA CHAIN.
OS Oncorhynchus keta (Chum salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8018;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93208376; Pubmed=1296820;
RA Coe I.R., Munro R., Sherwood N.M.;
RT "Isolation of different brain-specific isoforms of alpha-tubulins
RT from chum salmon (Oncorhynchus keta).";
RL DNA Seq. 3:257-262(1992).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROFIBULES. IT
CC BINDS TWO MOLES OF GTP. ONE AT AN EXCHANGEABLE SITE ON THE BETA
CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X66973; CAA47384.1; -
DR PIR; S25004; S25004.
DR InterPro; IPR000217; Tubulin.
DR InterPro; IPR003008; Tubulin_Fts2.
DR Pfam; PF00091; tubulin; 1.
DR PRINTS; PR01161; TUBULIN.
DR PRINTS; PR01162; ALPHATUBULIN.
DR PROSITE; PS00227; TUBULIN; 1.
KW Microtubules; GTP-binding; Multigene family.
FT NP_BIND 136 142 GTP (POTENTIAL).
FT SITE 444 444 INVOLVED IN POLYMERIZATION.
FT SITE 444 444 INVOLVED IN POLYMERIZATION.
SQ SEQUENCE 444 AA; 49314 MM; C2CD75432DD6BAF CRC64;

```

Query Match 1.6%; Score 7; DB 1; Length 444;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 13:05:17 ; Search time 77.39 Seconds

(Without alignments)
835,410 Million cell updates/sec

Title: US-09-787-083-4

Perfect score: 442
Sequence: 1 MKVSLSTLTLSTLPCFALLA.....YVHEATSPGVGLMDNMCL 442Scoring table:
OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 473505 seqs, 14627329 residues

Word size : 0

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL_17:*

- 1: sp.archaea:*
- 2: sp.bacteria:*
- 3: sp.fungi:*
- 4: sp.human:*
- 5: sp.invertebrate:*
- 6: sp.mammal:*
- 7: sp.mhc:*
- 8: sp.organelle:*
- 9: sp.phage:*
- 10: sp.plant:*
- 11: sp.podent:*
- 12: sp.virus:*
- 13: sp.vertebrate:*
- 14: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	2.0	382	2 09K0U7	09K0U7 neisseria m
2	9	2.0	409	2 09U721	09U721 neisseria m
3	9	2.0	862	2 09C162	09C162 pasteurilla
4	9	2.0	1432	10 09AX85	09AX85 oryza sativ
5	8	1.8	299	10 09M268	09M268 arabidopsis
6	8	1.8	306	2 09C122	09C122 pasteurilla
7	7	1.6	55	12 067608	067608 tomato gold
8	7	1.6	72	12 091RS5	091RS5 potato viru
9	7	1.6	72	12 091RS4	091RS4 potato viru
10	7	1.6	95	12 091RS3	091RS3 potato viru
11	7	1.6	96	12 091RS8	091RS8 potato viru
12	7	1.6	101	12 091RS6	091RS6 potato viru
13	7	1.6	122	2 09RU16	09RU16 delinococcus
14	7	1.6	124	2 091RS7	091RS7 potato viru
15	7	1.6	131	6 09GMD4	09GMD4 macaca fusc
16	7	1.6	138	12 091XQ7	091XQ7 potato viru
17	7	1.6	147	5 036035	036035 hexamita in
18	7	1.6	158	2 0929C1	0929C1 chlamydia p
19	7	1.6	169	2 0912A7	0912A7 streptomyce

20	7	1.6	186	10 09LXR9	09LXR9 arabidopsis
21	7	1.6	189	13 09DFT2	09DFT2 notochenia
22	7	1.6	198	4 09BUX9	09BUX9 homo sapien
23	7	1.6	212	13 091819	091819 ictalurus p
24	7	1.6	216	2 09X6H6	09X6H6 streptococc
25	7	1.6	218	2 09RF14	09RF14 streptococc
26	7	1.6	219	5 017594	017594 caenorhabdi
27	7	1.6	220	2 09XCX1	09XCX1 streptococc
28	7	1.6	220	5 036036	036036 spirocheteu
29	7	1.6	221	12 065002	065002 allia mo
30	7	1.6	227	5 09BMP4	09BMP4 glossina mo
31	7	1.6	233	6 09MZB0	09MZB0 ovis aries
32	7	1.6	236	13 09DRS8	09DRS8 notochenia
33	7	1.6	264	12 085260	085260 potato viru
34	7	1.6	265	9 038155	038155 bacterioph
35	7	1.6	267	9 099363	099363 bacterioph
36	7	1.6	267	12 085261	085261 potato viru
37	7	1.6	267	12 098628	098628 potato viru
38	7	1.6	267	12 086842	086842 potato viru
39	7	1.6	267	12 099026	099026 potato viru
40	7	1.6	267	12 005476	005476 potato viru
41	7	1.6	267	12 009711	009711 potato viru
42	7	1.6	267	12 09WH17	09WH17 potato viru
43	7	1.6	267	12 09QNM9	09QNM9 potato viru
44	7	1.6	267	12 09DUL0	09DUL0 potato viru
45	7	1.6	267	12 09DUX9	09DUX9 potato viru

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	382 AA.
09K0U7	09K0U7			
AC	09K0U7			
DT	01-OCT-2000 (TREMblrel. 15, Created)			
DT	01-OCT-2000 (TREMblrel. 15, Last sequence update)			
DT	01-JUN-2001 (TREMblrel. 17, Last annotation update)			
DE	PHOSPHOLIPASE A1, PUTATIVE.			
GN	NMB0464.			
OS	Neisseria meningitidis (serogroup B).			
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.			
OX	NCBI_TaxID=491;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MC58 / SEROGROUP B;			
RX	MEDLINE=2017575; PubMed=10710307;			
RA	Tetzelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,			
RA	Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,			
RA	Nelson W.C., Gwinn M.L., Deboy R., Peterson J.D., Hickey E.K.,			
RA	Halt D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,			
RA	Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,			
RA	Cotton M.D., Utterback T.R., Khouli H., Qin H., Yamathavan J.,			
RA	Gill J., Scarlato V., Maignani V., Piza M., Grandi G., Sun L.,			
RA	Smith H.O., Fraser C.M., Moxon E.R., Rappunli R., Venter J.C.;			
RT	"Complete genome sequence of Neisseria meningitidis serogroup B strain			
RT	MC58."			
RL	Science 287:1809-1815(2000).			
DR	EMBL: AE002403; AAF40901.1; -.			
DR	TIGR: NMB0464; -.			
DR	InterPro: IPR003187; PLAI.			
DR	Pfam: PF02253; PLAI; 1.			
KW	Complete proteome.			
SO	SEQUENCE 382 AA; 42714 MW; B46BA802F062E836 CRC64;			

Query Match	2.0%	Score 9:	DB 2:	Length 382;
Best Local Similarity	100.0%	Pred. No. 1.4;		
Matches 9;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 279 PEIFLTPV 287				
Db 221 PEIFLTPV 229				

```

RESULT 2
Q9JUT21
ID 09JUT21 PRELIMINARY: PRT: 409 AA.
AC 09JUT21:
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DE 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)
DE PUTATIVE PHOPHOLIPASE.
GN NMA2021.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBL_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Z2491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE-20222556; PubMed-10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holtroyd S.,
RA Jagers K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491."
RL Nature 404:502-506(2000).
DR EMBL: AL162757; CAB85240.1;
DR InterPro: IPR003187; Pfam:
DR Pfam: PF02253; Pfam: 1.
KW Complete proteome.
SQ SEQUENCE 409 AA; 45862 MW; CD6585B064D01A41 CRC64;

Query Match
Best Local Similarity 2.0%; Score 9; DB 2; Length 409;
Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIFLTPV 287
DB 248 PEIFLTPV 256

RESULT 3
Q9CL62
ID 09CL62 PRELIMINARY: PRT: 862 AA.
AC 09CL62;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)
DE CVGSI.
GN CVGSI OR PM1380.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBL_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PM70;
RX MEDLINE-21145866; PubMed-11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida PM70."
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
CC REGULATORY COMPONENTS OF SENSORY TRANSDUCTION SYSTEMS.
CC -1- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE
CC KINASES.
DR EMBL: AE006176; AAK03464.1;
DR InterPro: IPR000410; Bcrl_sens.
DR InterPro: IPR000658; DUF5.
DR InterPro: IPR003660; HAMF.
DR InterPro: IPR003594; HAMFase_C.
DR InterPro: IPR003661; His_kinA.

```

```

DR InterPro: IPR001789; Response_reg.
DR Pfam: PF00672; DUF5. 1.
DR Pfam: PF00072; response_reg. 1.
DR Pfam: PF00512; signal. 1.
DR PRINTS: PR00344; BCTRLSENSOR.
DR SMART: SM00304; HAMF. 1.
DR SMART: SM00387; HAMFase_C. 1.
DR SMART: SM00388; H1SKA. 1.
DR SMART: SM00448; RCC. 1.
KW Complete proteome; Kinase; Phosphorylation; Sensory transduction;
KW Transferase.
SQ SEQUENCE 862 AA; 97846 MW; AB6814A3B63626C0 CRC64;

Query Match
Best Local Similarity 2.0%; Score 9; DB 2; Length 862;
Pred. No. 2.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LSTLTSLT 13
DB 293 LSTLTSLT 301

RESULT 4
Q9AX85
ID 09AX85 PRELIMINARY: PRT: 1432 AA.
AC 09AX85;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)
DE PUTATIVE ABC TRANSPORTER PROTEIN.
GN P0410E03.7.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBL_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone: P0410E03."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP002844; BAB21276.1;
SQ SEQUENCE 1432 AA; 160234 MW; 04E8CF2CA08D993 CRC64;

Query Match
Best Local Similarity 2.0%; Score 9; DB 10; Length 1432;
Pred. No. 4.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 FTPLSLSF 178
DB 830 FTPLSLSF 838

RESULT 5
Q9M268
ID 09M268 PRELIMINARY: PRT: 299 AA.
AC 09M268;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)
DE RAV-LIKE PROTEIN.
GN P21P14.140.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBL_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.

```



```

RESULT 7
067608
ID 067608 PRELIMINARY; PRT; 55 AA.
AC 067608;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE SUGENOMIC DNA DERIVED FROM DNA B CCDS - COVALENTLY CLOSED CIRCULAR
DE DOUBLE-STRANDED MOLECULE.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OC NCBI_TaxID=10831;
RN [1]
RP SEQUENCE FROM N.A.

```

RESULT	9		
091ZS4			
ID	091ZS4	PRELIMINARY;	PRT; 72 AA.
AC	091ZS4;		
DT	01-OCT-2000 (TREMBLrel. 15, Created)		
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)		
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)		
DE	COAT PROTEIN (FRAGMENT).		
GN	CP.		
OS	Potato virus Y strain NTN (PVY(NTN)).		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;		
CC	Potyvirus.		
OX	NCBI_TaxID=122280;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	SRAIN-TUBER NECROSING (NTN);		
RC	Cervoska N., Filigarova M., Moravec T., Petrzlik K.;		

RT "Differences in nucleotide and amino acid sequences of N-terminal
RT parts of coat proteins among isolates of potato virus Y-NTN strain."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF228634; AAF67875.1; -
FT NON_TER 1
FT SEQUENCE 72 AA; 7692 MW; 4F129B870523B482 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 72;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKQ 162
Db 11 TKKDAKQ 17

RESULT 10
Q91ZS3 PRELIMINARY; PRT; 95 AA.
AC Q91ZS3;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y strain NTN (PVY(NTN)).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Polyvirus.
OX NCBI_TaxID=122280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUBER NECROSING (NTN);
RA Cerovska N., Filigayova M., Moravec T., Petrizik K.;
RT "Differences in nucleotide and amino acid sequences of N-terminal
RT parts of coat proteins among isolates of potato virus Y-NTN strain."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF228635; AAF67876.1; -
FT NON_TER 1
FT SEQUENCE 95 AA; 10446 MW; 28C06891777C9928 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 95;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKQ 162
Db 11 TKKDAKQ 17

RESULT 11
Q91ZS8 PRELIMINARY; PRT; 96 AA.
AC Q91ZS8;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y (strain N) (PVY).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Polyvirus.
OX NCBI_TaxID=12219;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NECROSING (N);
RA Cerovska N., Filigayova M., Moravec T., Petrizik K.;
RT "Differences in nucleotide and amino acid sequences of N-terminal
RT parts of coat proteins among isolates of potato virus Y-NTN strain."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF228630; AAF67871.1; -
FT NON_TER 1
FT SEQUENCE 96 AA; 10559 MW; 89C3A5675B9784CA CRC64;

Query Match 1.6%; Score 7; DB 12; Length 96;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKQ 162
Db 11 TKKDAKQ 17

RESULT 12
Q91ZS6 PRELIMINARY; PRT; 101 AA.
AC Q91ZS6;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE COAT PROTEIN (FRAGMENT).
GN CP.
OS Potato virus Y strain NTN (PVY(NTN)).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Polyvirus.
OX NCBI_TaxID=122280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUBER NECROSING (NTN);
RA Cerovska N., Filigayova M., Moravec T., Petrizik K.;
RT "Differences in nucleotide and amino acid sequences of N-terminal
RT parts of coat proteins among isolates of potato virus Y-NTN strain."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF228632; AAF67873.1; -
FT NON_TER 1
FT SEQUENCE 101 AA; 11112 MW; B5B5F51C3D5FE9BD CRC64;

Query Match 1.6%; Score 7; DB 12; Length 101;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TKKDAKQ 162
Db 11 TKKDAKQ 17

RESULT 13
Q9RU16 PRELIMINARY; PRT; 122 AA.
AC Q9RU16;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE HYPOTHETICAL 13.7 KDA PROTEIN.
GN DR1580.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RI;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Uitterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Milton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;

RT "genome sequence of the radioresistant bacterium *Deinococcus*
 RL *radiodurans* RI.";
 DR EMBL: AE002002; AAF1150.1; -.
 DR TIGR: DR1580; -.
 DR InterPro: IPR000325; Glyoxalase_1.
 DR Pfam: PF00903; Glyoxalase; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 122 AA; 13683 MW; 7BA836DCAD786E51 CRC64;

Query Match 1.6%; Score 7; DB 2; Length 122;
 Best Local Similarity 100.0%; Pred. No. 58;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 EHEPELY 76
 Db 115 EHEPELY 121

RESULT 14
 Q912S7 PRELIMINARY; PRT; 124 AA.
 AC Q912S7;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
 DE COAT PROTEIN (FRAGMENT).
 GN CP.
 OS Potato virus Y strain NTN (PVY(NTN)).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 OC Potyvirus;
 OX NCBI_TaxID=122280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TUBER NECROSING (NTN);
 RA Czerovska N., Filigayova M., Moravec T., Petrlik K.;
 RT "Differences in nucleotide and amino acid sequences of N-terminal
 parts of coat proteins among isolates of potato virus Y-NTN strain.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF228631; AAF67872.1; -.
 FT NON_TER 1 1
 FT NON_TER 124 124
 SQ SEQUENCE 124 AA; 13654 MW; DE3618D620BA7080 CRC64;

Query Match 1.6%; Score 7; DB 12; Length 124;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 156 TKKDAQ 162
 Db 11 TKKDAQ 17

RESULT 15
 Q9GMD4 PRELIMINARY; PRT; 131 AA.
 AC Q9GMD4;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE ALPHA-TUBULIN (FRAGMENT).
 OS Macaca fuscata (Japanese macaque).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9542;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99270338; PubMed=10340513;
 RA Okuno H., Tokuyama W., Li Y.X., Hashimoto T., Miyashita Y.;
 RT "Quantitative evaluation of neurotrophin and trk mRNA expression in

RT visual and limbic areas along the occipito-temporo-hippocampal pathway
 RT in adult macaque monkeys.";
 RL J. Comp. Neurol. 408:378-398(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Hashimoto T., Tokuyama W., Okuno H., Li Y., Miyashita Y.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF286378; AAG00533.1; -.
 DR InterPro: IPR03008; tubulin_FtsZ.
 DR Pfam: PF00091; tubulin; 1.
 KW GTP-binding.
 KW GTP-binding.
 FT NON_TER 1 1
 FT NON_TER 131 131
 SQ SEQUENCE 131 AA; 14575 MW; A07D9C90B0641FCE CRC64;

Query Match 1.6%; Score 7; DB 6; Length 131;
 Best Local Similarity 100.0%; Pred. No. 62;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 IKTKRSI 120
 Db 110 IKTKRSI 116

Search completed: May 9, 2002, 13:05:18
 Job time: 613 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:54:58 ; Search time 72.68 Seconds

(without alignments)
450,473 Million cell updates/sec

Title: US-09-787-083-6

Sequence: 1 MKVSLSTLTSLTSCPAALLA.....YHNEATSFVGIMLNDMGL 442

Scoring table:

OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 522463 seqs, 74073290 residues

Word size : 0

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: /SIDS8/gcgdata/geneseq/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	442	100.0	442	21	AAV85270
2	390	88.2	442	21	AAV85269
3	316	71.5	442	21	AAV85271
4	224	50.7	442	21	AAV85268
5	9	2.0	370	21	AAV75156
6	9	2.0	370	21	AAV75157
7	9	2.0	370	21	AAV75158
8	9	2.0	374	21	AAV70629
9	9	2.0	375	21	AAV70628
10	8	1.8	278	21	AAV1137
11	8	1.8	283	21	AAV1136

12	8	1.8	299	21	AAV1135
13	7	1.6	16	15	AAV52127
14	7	1.6	33	22	AAV76069
15	7	1.6	50	21	AAV79483
16	7	1.6	64	21	AAV60322
17	7	1.6	82	21	AAV01488
18	7	1.6	108	21	AAV37811
19	7	1.6	116	21	AAV7313
20	7	1.6	136	21	AAV37810
21	7	1.6	157	21	AAV29867
22	7	1.6	169	20	AAV34661
23	7	1.6	181	21	AAV37809
24	7	1.6	268	13	AAV21542
25	7	1.6	302	13	AAV24254
26	7	1.6	316	21	AAV39500
27	7	1.6	324	21	AAV39499
28	7	1.6	326	13	AAV24356
29	7	1.6	365	21	AAV39498
30	7	1.6	381	20	AAV60241
31	7	1.6	448	21	AAV191960
32	7	1.6	523	21	AAV44285
33	7	1.6	552	16	AAV67007
34	7	1.6	666	20	AAV21871
35	7	1.6	683	20	AAV21869
36	7	1.6	684	20	AAV21875
37	7	1.6	715	20	AAV21860
38	7	1.6	715	20	AAV21860
39	7	1.6	900	20	AAV6264
40	7	1.6	934	17	AAV07637
41	7	1.6	1026	17	AAV97374
42	7	1.6	1026	20	AAV94678
43	7	1.6	1026	20	AAV92358
44	7	1.6	1026	22	AAV35154
45	7	1.6	1026	22	AAV35154

ALIGNMENTS

RESULT 1
ID AAV85270 standard; Protein: 442 AA.
AC AAV85270;
XX
DT 29-JUN-2000 (first entry)
XX
DE BASB034 amino acid sequence #3.
XX
KW Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KW vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW hearing loss; antibacterial drug.
XX
OS Moraxella catarrhalis.
PN WO200015802-A1.
PD 23-MAR-2000.
XX
PE 14-SEP-1999; 99WO-EP06781.
XX
PR 14-SEP-1998; 98GB-0020002.
PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Ruelle J;
XX
DR WPI, 2000-271440/23.
DR N-PSDB; AAA10702.
XX
PT Novel BASB034 polynucleotides and polypeptides from Moraxella
catarrhalis used to prepare vaccines against bacterial infections

XX Claim 3; Page 68; 106pp; English.
XX
CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC strain Mc2913. The invention relates to BASB034 polypeptides from
CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
CC polynucleotides and polypeptides may be employed as research reagents and
CC material for the discovery of treatments and diagnostics for diseases,
CC particularly human diseases. They are particularly used to diagnose and
CC treat M. catarrhalis infections. They can be particularly used for diagnosis of
CC disease, staging of disease, or determining response of an infectious
CC organism to drugs. The polynucleotides may be used as a source for
CC hybridization probes, and for screening of genetic mutations, serotype,
CC organism or strain identification, identification of mutations in BASB034
CC sequences, and as components of arrays which are useful for diagnostic
CC and prognostic purposes. The polypeptides can be used to produce
CC antibodies. The polypeptides can also be used in vaccine formulations,
CC and to identify agonists and antagonists. The polypeptides, antibodies,
CC agonists and antagonists (which are bacteriostatic) are used for the
CC treatment and prevention of diseases such as otitis media in infants and
CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
CC invasive diseases, and chronic otitis media with hearing loss. The
CC polypeptides, agonists and antagonists are also used for screening of
CC antibacterial drugs. The BASB034 products of the invention can be used
CC screen for new antibacterial compounds that may target resistant
CC bacteria.
XX
XX Sequence 442 AA;

Query Match 100.0%; Score 442; DB 21; Length 442;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKSLSTLTSLTSLSCFAIIAQAKAVNPVAFVDEVRSENDLGDDELPIDVQASATQSA 60
DB 1 mkvslstltsltslscfaiiaqakavnpvafvdevrseendlgdelpidvqsaqsa 60
QY 61 STDTANPLDEHEPELYTTALENKTMLINCALNODIMRLACYDTLVHGETPAVITKTRST 120
DB 61 stdtanpldehepe lyttalenktmlincalnodimrlacydtlvhgetpavltktrsl 120
QY 121 RLDETTIMQTIKGRPOVYVVOETDPIFLMGNEKGMLTKKDAKOLEYAKQTPISLSDLD 180
DB 121 rldettimqti kgrpovyvvoetdpi flmgnekgmtkkdakeleyakqtpislslsdld 180
QY 181 RNNTPPLMSRRPHNPMYVLPITFMHGKRPNSPNTFSEHARQTPNFEFAPELKFOYSVYKA 240
DB 181 rnntpplmsrrphn pmyvlpitfmhgk rpnspntfseharqtpnfe fapelkfoysvyka 240
QY 241 AEDLMCTSDLMFGYTOQSHMQIFNCKNSRPFVNDYQPEIFLTQPVYSDLPMDGKVRMT 300
DB 241 aedlmctsdlmf gytqoshmqifnc knsrpfvndyqpeifltqpvysdlpmdgkvrm 300
QY 301 GMGAVHHSNGESAKLSNSNRATILMAGMEKNLTVMPTINGRTFKESGSGQPDNDPIILD 360
DB 301 gmgavhhsngesa klsnswratilmag meknltvmptingrtfkessgsgqpdndpild 360
QY 361 YYGCGDVRFYQLEKNSNIGTVRYNPRSGKALQLDYVYPLGKGISGYQIQFOGYQOSTL 420
DB 361 yygcvdrf yqleknsgnigtvry nprsgkalqldyvyp lggkigsgyqi qfoqyqostl 420
QY 421 IDYNHEATSGVGLMDMNGML 442
DB 421 idynheatsg vglmdmngml 442

RESULT 2
AA185269
ID AA185269 standard; Protein; 442 AA.
XX
AC AA185269;
XX

DT 29-JUN-2000 (first entry)
XX
XX BASB034 amino acid sequence #2.
DE
XX Moraxella catarrhalis infection: BASB034; diagnosis: staging;
KW vaccine: bacteriostatic; treatment: prevention; otitis media; pneumonia;
KW sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KW hearing loss; antibacterial drug.
XX
OS Moraxella catarrhalis.
XX
PN WO200015802-A1.
XX
PD 23-MAR-2000.
XX
PF 14-SEP-1999; 99WO-EP06781.
XX
PR 14-SEP-1998; 98GB-0020002.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Ruelle J;
XX
PI
XX
DR WPI: 2000-271440/23.
XX
DR N-PsDB; AAA10701.
XX
PT Novel BASB034 polynucleotides and polypeptides from Moraxella
PT catarrhalis used to prepare vaccines against bacterial infections
XX
XX
PS Claim 3; Page 67; 106pp; English.

CC This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC strain Mc2908. The invention relates to BASB034 polypeptides from
CC M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
CC polynucleotides and polypeptides may be employed as research reagents and
CC material for the discovery of treatments and diagnostics for diseases,
CC particularly human diseases. They are particularly used to diagnose and
CC treat M. catarrhalis infections. They can be used for diagnosis of
CC disease, staging of disease, or determining response of an infectious
CC organism to drugs. The polynucleotides may be used as a source for
CC hybridization probes, and for screening of genetic mutations, serotype,
CC organism or strain identification, identification of mutations in BASB034
CC sequences, and as components of arrays which are useful for diagnostic
CC and prognostic purposes. The polypeptides can be used to produce
CC antibodies. The polypeptides can also be used in vaccine formulations,
CC and to identify agonists and antagonists. The polypeptides, antibodies,
CC agonists and antagonists (which are bacteriostatic) are used for the
CC treatment and prevention of diseases such as otitis media in infants and
CC children, pneumonia in elderlies, sinusitis, nosocomial infections and
CC invasive diseases, and chronic otitis media with hearing loss. The
CC polypeptides, agonists and antagonists are also used for screening of
CC antibacterial drugs. The BASB034 products of the invention can be used
CC screen for new antibacterial compounds that may target resistant
CC bacteria.
XX
XX Sequence 442 AA;

Query Match 88.2%; Score 390; DB 21; Length 442;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 VQASATOSASTDTANPLDEHEPELYTTALENKTMLINCALNODIMRLACYDTLVHGETPA 112
DB 53 vqatsqasastd tanpldehepe lyttalenktmlincaln odimrlacydtlvhgetp 112
QY 113 VIKTKSRIDETIMQTIKGRPOVYVVOETDPIFLMGNEKGMLTKKDAKOLEYAAQOFPP 172
DB 113 viktksrirdet imqti kgrpovyvvoetdpi flmgnekgmtkkdakeleyaakqfcp 172
QY 173 LSLSPDLDRNNTPPLMSRRPHNPMYVLPITFMHGKRPNSPNTFSEHARQTPNFEFAPELK 232
DB 173 lslspdl drnn t p l m s s r p h n p m y v l p i t f m h g k r p n s p n t f s e h a r q t p n f e f a p e l k f 232

QY 233 QVSVKVAEDLMCTSDDLMEFGTYQOSHWOJFNCKNSRPFRVNDYQPELITQPVYSDLP 292
CC |
CC |
CC |
Db 233 qvsvkvaedlmctsdldwlfgyfcgshwqjfncknsrfrvndyqpeflitcpvydldp 292
QY 293 WDGVKRMIGGAVHNSGESAKLSRSMNRAVYIMAGMEKNTVMPRIWGRFKEGSGSQP 352
CC |
CC |
CC |
Db 293 wdgvkrmigmavhnsngesaklsrsmnraylmagmeknltvmpriwgrlfksgsgsqp 352
QY 353 DDNPDILDYGYGDVRFYQLEENKSNISGTVRNPBSGKGMQLDYYVYPLGKISGYEQI 412
CC |
CC |
CC |
Db 353 ddnpdildygygdvrflyqleemknsisgtrvnpbsgkqaldyvyplgkisyfqi 412
QY 413 FQGTGQSLIDYNNHATSFVGLMLNDMAGL 442
CC |
CC |
CC |
Db 413 fggygslldynheatsfygvlmndwmg1 442

RESULT 3
AAV85271
ID AAV85271 standard; Protein: 442 AA.
XX
XX AAV85271;
XX
XX 29-JUN-2000 (first entry)
XX
XX BASB034 amino acid sequence #4.
DE
XX
XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KM vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KM sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KM hearing loss; antibacterial drug.
XX
XX
XX Moraxella catarrhalis.
OS
XX
XX WC200015802-A1.
PN
XX
XX 23-MAR-2000.
PD
XX
XX 14-SEP-1999; 99WC-EP06781.
PF
XX
XX 14-SEP-1998; 98GB-0020002.
PR
XX
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
PA
XX
XX Ruelle J;
PI
XX
XX WPI: 2000-271440/23.
DR
XX
XX N-PSDB: AAA10703.

Novel BASB034 polynucleotides and polypeptides from Moraxella
catarrhalis used to prepare vaccines against bacterial infections -
PS
XX
XX Claim 3; Page 69; 106pp; English.

This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
strain Mc2969. The invention relates to BASB034 polypeptides from
M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The BASB034
polynucleotides and polypeptides may be employed as research reagents and
material for the discovery of treatments and diagnostics for diseases,
CC particularly human diseases. They are particularly used to diagnose and
CC treat M. catarrhalis infections. They can be used for diagnosis of
CC disease, staging of disease, or determining response of an infectious
CC organism to drugs. The polynucleotides may be used as a source for
CC hybridization probes, and for screening of genetic mutations, serotype,
CC organism or strain identification, identification of mutations in BASB034
CC sequences, and as components of arrays which are useful for diagnostic
CC and prognostic purposes. The polypeptides can be used to produce
CC antibodies. The polypeptides can also be used in vaccine formulations,
CC and to identify agonists and antagonists. The polypeptides, antibodies,
CC agonists and antagonists (which are bacteriostatic) are used for the
CC treatment and prevention of diseases such as otitis media in infants and
CC children, pneumonia in elderlies, sinusitis, nosocomial infections and

CC invasive diseases, and chronic otitis media with hearing loss. The
CC polypeptides, agonists and antagonists are also used for screening of
CC antibacterial drugs. The BASB034 products of the invention can be used
CC screen for new antibacterial compounds that may target resistant
CC bacteria.
XX
XX
SQ Sequence 442 AA;
Query Match 71.5%; Score 316; DB 21; Length 442;
Best Local Similarity 99.8%; Pred. No. 3.6e-298;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 26 AVPNPVAFVDEVRSENDLGQDNELPIDYQASVTSIDTANPDEHEPELTTALENKT 85
CC |
CC |
CC |
Db 26 avpnpvafvdevrsendlgdneipidvqsatgsastdantpdehepeyltalenk 85
QY 86 LINCASALNODIMRLACYDTLVHGFEPVAVIKTKRSIRIDETIMOTIKGPOVYQETDP 145
CC |
CC |
CC |
Db 86 lincsalnqdimrlacydclvhgefpaviktkrsiridetlwtlqkqpvvyqetdcp 145
QY 146 FLMGNEKGMILTKKDAKOLEYAAKQFTPLSLSFDDRNTPLMSSRPHPNPTVLPFMH 205
CC |
CC |
CC |
Db 146 flmgnekgmiltkkdkakqleyakqftplsfsfdrrntplmssrphnptvlpf 205
QY 206 PNRSPNTPSHFARQFTPNFPAPELKFOVSVKVAEDLMCTSDDLMEFGTYQOSHWOJ 265
CC |
CC |
CC |
Db 206 pnrspntpshfarqftpnfpaapelkfovsyvkvaedlmctsdldwlfgyfcgshw 265
QY 266 GKNSRPPFVHXYOPELITQPVYSDLPMDGKVRMIGGAVHNSGESAKLSRSMNRAV 325
CC |
CC |
CC |
Db 266 gknsrpfvhyqpeflitcpvydldpmdgkvrnmigmavhnsngesaklsrsmnray 325
QY 326 AGMEKMLTVMPRIWGRIFKEGSGSQPDNDPILDYGYGDVRFYQLEENKSNISGTV 385
CC |
CC |
CC |
Db 326 agmekmltvmprlwgriwgrifksgsgsqpdndpildygygdvrflyqleemkns 385
QY 386 NPSGKGMQLDYYVYPLGKISGYFQGTGQSLIDYNNHATSFVGLMLNDMAGL 442
CC |
CC |
CC |
Db 386 npsgkqaldyvyplgkisyfqi fggygslldynheatsfygvlmndwmg1 442

RESULT 4
AAV85268
ID AAV85268 standard; Protein: 442 AA.
XX
XX
XX AAV85268;
XX
XX
XX 29-JUN-2000 (first entry)
XX
XX
XX BASB034 amino acid sequence #1.
DE
XX
XX Moraxella catarrhalis infection; BASB034; diagnosis; staging;
KM vaccine; bacteriostatic; treatment; prevention; otitis media; pneumonia;
KM sinusitis; nosocomial infection; invasive disease; chronic otitis media;
KM hearing loss; antibacterial drug.
XX
XX
XX Moraxella catarrhalis.
OS
XX
XX WC200015802-A1.
PN
XX
XX 23-MAR-2000.
PD
XX
XX 14-SEP-1999; 99WC-EP06781.
PF
XX
XX 14-SEP-1998; 98GB-0020002.
PR
XX
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
PA
XX
XX Ruelle J;
PI
XX
XX WPI: 2000-271440/23.
DR
XX
XX N-PSDB: AAA10700.

```

XX  Novel BASB034 polynucleotides and polypeptides from Moraxella
PT  catarrhalis used to prepare vaccines against bacterial infections
XX
XX  Claim 3; Fig 2; 106pp; English.
XX
CC  This sequence represents a Moraxella catarrhalis BASB034 polypeptide from
CC  strain Mc2931 (ATCC 43617). The invention relates to BASB034 polypeptides
CC  from M. catarrhalis strains Mc2931, Mc2908, Mc2913 and Mc2969. The
CC  BASB034 polynucleotides and polypeptides may be employed as research
CC  reagents and material for the discovery of treatments and diagnostics for
CC  diseases, particularly human diseases. They are particularly used to
CC  diagnose and treat M. catarrhalis infections. They can be used for
CC  diagnosis of disease, staging of disease, or determining response of an
CC  infectious organism to drugs. The polynucleotides may be used as a source
CC  for hybridization probes, and for screening of genetic mutations.
CC  serotype, organism or strain identification, identification of mutations
CC  in BASB034 sequences, and as components of arrays which are useful for
CC  diagnostic and prognostic purposes. The polypeptides can be used to
CC  produce antibodies. The polypeptides can also be used in vaccine
CC  formulations, and to identify agonists and antagonists. The polypeptides,
CC  antibodies, agonists and antagonists (which are bacteriostatic) are used
CC  for the treatment and prevention of diseases such as otitis media in
CC  infants and children, pneumonia in elders, sinusitis, nosocomial
CC  infections and invasive diseases, and chronic otitis media with hearing
CC  loss. The polypeptides, agonists and antagonists are also used for
CC  screening of antibacterial drugs. The BASB034 products of the invention
CC  can be used screen for new antibacterial compounds that may target
CC  resistant bacteria.
XX
SQ  Sequence 442 AA;

Query Match 50.7%; Score 224; DB 21; Length 442;
Best Local Similarity 100.0%; Pred. No. 6.6e-209; Indels 0; Gaps 0;
Matches 224; Conservative 0; Mismatches 0;

QY 219 QTPNFRAPDLRFQVSVKVKAEDLMGTSDLMFGYTQSHQIFNGKNSRPERVHDYQ 278
    |||||||
DB 219 qfipnefrapelkfyavkvkaeedlwtidsdlwfytqshwqifngksrperfvhdqy 278

QY 279 PEIFLTQPVSDLPMDCKVMKMGAVHHSNGESAKTSRSMNRYLMAAGHEMKRLTYMPR 338
    |||||||
DB 279 peliftqpvysdlpwgdkvrmigmavhshngesaklsrsmnrylmaaghewmkrltymp 338

QY 339 IMGRIFEGSGSQPDNDPDLIDYGYGDAVRELYOLENKSINSGTVRYNPRSGKALOLDY 398
    |||||||
DB 339 lmgriifegsgsqpdndpdlidyygygdvrflyqjenksnsgtvtvrynprrsgkaloldy 398

QY 399 VYPLGKIGISGYFOIFOGYGSLIDYVNEHATSPFGVGLMLNDMGMGL 442
    |||||||
DB 399 vylpgkigisgyfqiifgygslidynheatsfvgvglmndwmgmgl 442

RESULT 5
AAV75156
ID AAV75156 standard; Protein: 370 AA.
XX
AC AAV75156;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria gonorrhoeae ORF 582 protein sequence SEQ ID NO:1786.
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KM antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
KW antibacterial; gene therapy.
XX
XX Neisseria meningitidis.
OS Neisseria gonorrhoeae.
XX
XX MO9957280-AZ.
XX
XX 11-NOV-1999.
PD

```

```

XX 30-APR-1999; 99WO-US09346.
XX
XX 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
XX (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalzo E, Scarselli M;
PI Tettelin H, Venter JC;
XX
DR WPI; 2000-062150/05.
DR N-PsDB; AA253918.
XX
PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics
XX
PS Claim 2; Page 903; 1453pp; English.
XX
CC AA253015 to AA254536, AA254577 to AA254615, and AAV74253 to AAV75941
CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
CC and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
SQ Sequence 370 AA;

Query Match 2.0%; Score 9; DB 21; Length 370;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 PEIFLTQPV 287
    |||||||
DB 209 peliftqpv 217

RESULT 6
AAV75157
ID AAV75157 standard; Protein: 370 AA.
XX
AC AAV75157;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 582 protein sequence SEQ ID NO:1788.
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KM antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
KW antibacterial; gene therapy.
XX
XX Neisseria meningitidis.
OS Neisseria meningitidis.
XX
XX MO9957280-AZ.
XX
XX 11-NOV-1999.
PD

```


PF 30-APR-1999; 99WO-US09346.
 XX
 PR 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX
 PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX
 DR WPI: 2000-062150/05.
 DR N-PDB: AA253919.
 XX
 PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics
 XX
 PS Claim 2; Page 904; 1453pp; English.
 XX
 CC AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941
 CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA255473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 CC
 XX
 SQ Sequence 370 AA;

Query Match 2.0%; Score 9; DB 21; Length 370;
 Best Local Similarity 100.0%; Pred. No. 2.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 279 PEIFLTQPV 287
 |||||||||
 Db 209 pelfltqpv 217

RESULT 7
 AAY75158
 ID AAY75158 standard; Protein; 370 AA.
 XX
 AC AAY75158;
 XX
 DT 21-MAR-2000 (first entry)
 XX
 DE Neisseria meningitidis ORF 582 protein sequence SEQ ID NO:1790.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
 KW antibacterial; gene therapy.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO957280-A2.
 XX
 PD 11-NOV-1999.
 XX
 PF 30-APR-1999; 99WO-US09346.

XX
 PR 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX
 PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX
 DR WPI: 2000-062150/05.
 DR N-PDB: AA253920.
 XX
 PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics
 XX
 PS Claim 2; Page 905; 1453pp; English.
 XX
 CC AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941
 CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA255473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 CC
 XX
 SQ Sequence 370 AA;

Query Match 2.0%; Score 9; DB 21; Length 370;
 Best Local Similarity 100.0%; Pred. No. 2.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 279 PEIFLTQPV 287
 |||||||||
 Db 209 pelfltqpv 217

RESULT 8
 AAY70629
 ID AAY70629 standard; Protein; 374 AA.
 XX
 AC AAY70629;
 XX
 DT 18-JUL-2000 (first entry)
 XX
 DE Neisseria meningitidis serogroup B strain H44/76 BASB033 protein.
 XX
 KW BASB033; diagnosis; prophylaxis; treatment; antibacterial; vaccine;
 KW Neisseria meningitidis infection.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO200015801-A1.
 XX
 PD 23-MAR-2000.
 XX
 PF 09-SEP-1999; 99WO-EP06718.
 XX
 PR 14-SEP-1998; 98GB-0020003.

```
XX (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.
PA
XX
XX Ruelle J;
PI
XX
DR WPI: 2000-271439/23.
DR N-PSDB; AA52134.
XX
XX Isolated BASB033 polypeptides and polynucleotides of Neisseria
PT meningitidis, useful for diagnosis, prophylaxis and treatment of N.
PT meningitidis infection -
XX
PS Claim 4; Page 59; 93pp; English.
XX
XX The present sequence is a BASB033 protein from
CC Neisseria meningitidis serogroup B strain H44/76. The protein
CC shows homology to the Klebsiella pneumoniae outer membrane
CC phospholipase A. The present sequence is useful for diagnosis,
CC prophylaxis and treatment of N. meningitidis infection. It may also be
CC used for the discovery and development of antibacterial compounds and
CC in vaccine compositions.
XX
SQ Sequence 374 AA;

Query Match 2.0%; Score 9; DB 21; Length 374;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 PEIFLTQPV 287
Db 213 pelfltqpv 221

RESULT 9
AAY70628
ID AAY70628 standard; Protein; 375 AA.
XX
XX AAY70628;
AC
XX
DT 18-JUL-2000 (first entry)
XX
DE Neisseria meningitidis serogroup B strain ATCC13090 BASB033 protein.
XX
XX BASB033; diagnosis; prophylaxis; treatment; antibacterial; vaccine;
KM Neisseria meningitidis infection.
XX
XX Neisseria meningitidis.
OS
XX
XX WO200015801-A1.
FN
XX
XX 23-MAR-2000.
PD
XX
PF 09-SEP-1999; 99WO-EP06718.
XX
PR 14-SEP-1998; 98GB-0020003.
XX
XX (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.
PA
XX Ruelle J;
PI
XX
DR WPI: 2000-271439/23.
DR N-PSDB; AA52133.
XX
XX Isolated BASB033 polypeptides and polynucleotides of Neisseria
PT meningitidis, useful for diagnosis, prophylaxis and treatment of N.
PT meningitidis infection -
XX
XX Claim 4; Page 58; 93pp; English.
XX
XX The present sequence is a BASB033 protein from
CC Neisseria meningitidis serogroup B strain ATCC13090. The protein
CC shows homology to the Klebsiella pneumoniae outer membrane
```

```
CC phospholipase A. The present sequence is useful for diagnosis,
CC prophylaxis and treatment of N. meningitidis infection. It may also be
CC used for the discovery and development of antibacterial compounds and
CC in vaccine compositions.
XX
SQ Sequence 375 AA;

Query Match 2.0%; Score 9; DB 21; Length 375;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 PEIFLTQPV 287
Db 214 pelfltqpv 222

RESULT 10
AAG1137
ID AAG1137 standard; Protein; 278 AA.
XX
XX AAG1137;
AC
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 9741.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX 25-FEB-1999; 99US-0121825.
XX
XX 05-MAR-1999; 99US-0123180.
XX
XX 09-MAR-1999; 99US-0123548.
XX
XX 23-MAR-1999; 99US-0125788.
XX
XX 25-MAR-1999; 99US-0126264.
XX
XX 29-MAR-1999; 99US-0126785.
XX
XX 01-APR-1999; 99US-0127462.
XX
XX 06-APR-1999; 99US-0128234.
XX
XX 08-APR-1999; 99US-0128714.
XX
XX 16-APR-1999; 99US-0129845.
XX
XX 19-APR-1999; 99US-0130077.
XX
XX 21-APR-1999; 99US-0130449.
XX
XX 23-APR-1999; 99US-0130510.
XX
XX 28-APR-1999; 99US-0130891.
XX
XX 30-APR-1999; 99US-0131449.
XX
XX 30-APR-1999; 99US-0132048.
XX
XX 30-APR-1999; 99US-0132407.
XX
XX 04-MAY-1999; 99US-0132484.
XX
XX 05-MAY-1999; 99US-0132485.
XX
XX 06-MAY-1999; 99US-0132486.
XX
XX 06-MAY-1999; 99US-0132487.
XX
XX 07-MAY-1999; 99US-0132863.
XX
XX 11-MAY-1999; 99US-0134256.
XX
XX 14-MAY-1999; 99US-0134218.
XX
XX 14-MAY-1999; 99US-0134219.
XX
XX 14-MAY-1999; 99US-0134221.
XX
XX 14-MAY-1999; 99US-0134370.
XX
XX 18-MAY-1999; 99US-0134768.
XX
XX 19-MAY-1999; 99US-0134941.
XX
XX 20-MAY-1999; 99US-0135124.
XX
XX 21-MAY-1999; 99US-0135353.
XX
XX 24-MAY-1999; 99US-0135629.
XX
XX 25-MAY-1999; 99US-0136021.
XX
XX 27-MAY-1999; 99US-0136392.
```

PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158222.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159350.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159658.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.88; Score 8; DB 21; Length 283;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 174 SLSEFLDR 181
|||||||

Db 276 slsfdldr 283
RESULT 12
AC AG1135
ID AG1135 standard; Protein; 299 AA.
XX
AC AG1135;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 9739.
XX
KW Protein identification: signal transduction pathway; metabolic pathway;
KW hydrolisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127452.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137232.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139763.
PR 18-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141642.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150567.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151530.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158332.
PR 11-OCT-1999; 99US-0158369.
PR 12-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160778.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161923.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.88; Score 8; DB 21; Length 299;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 SLSFDLDR 181
Db 292 SLSFDLDR 299

RESULT 13
AAK52127
ID AAK52127 standard; Peptide; 16 AA.
XX

AC AAR52127;
 XX
 DT 27-SEP-1996 (first entry)
 XX
 DE Mouse light chain surface patch S03410.
 XX
 DE antibody: humanised; murine; human; heavy chain; light; variable;
 KW framework region; complementarity determining region; reshaping;
 KM modelling; surface residue; modify.
 XX
 OS Mus sp.
 XX
 PN EP592106-A1.
 XX
 PD 13-APR-1994.
 XX
 PF 07-SEP-1993; 93EP-0307051.
 XX
 PR 09-SEP-1992; 92US-0942245.
 XX
 PA (PEDE/) PEDERSEN J T.
 PA (IMMU-) IMMUNOGEN INC.
 XX
 PI Guild BC, Pedersen JT, Rees AR, Roguska MA, Searle SMJ;
 XX
 DR WPI; 1994-120230/15.
 XX
 PT Method of resurfacing of rodent antibodies to produce humanised
 PT antibody forms - for producing non-human antibodies with improved
 PT therapeutic efficiency by presenting human surface on V-region
 XX
 PS Example 1; Page 14; 230pp; English.
 XX
 CC Modification of a rodent antibody or fragment by resurfacing in order
 CC to produce a humanised rodent antibody can be determined by calculating
 CC homology between murine and human antibody surfaces. In order to test
 CC the resurfacing approach of the invention, three humanisation
 CC experiments were set up. (1) traditional loop grafting; (2) resurfacing
 CC approach using most similar chain; and (3) resurfacing approach using
 CC human sequences with most similar surface residues. AAR52069-159 are the
 CC surface residue patterns in mouse light chain antibody variable regions.
 CC These "patches" were used in the third method, where rodent light and
 CC heavy chains were matched and the most similar human sequence found
 CC independently only over the surface residues indicated in AAR52030-67.
 XX
 SQ Sequence 16 AA:
 Query Match 1.6%; Score 7; DB 15; Length 16;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 387 PRSGKA 393
 Db 4 prsgka 10
 RESULT 14
 AAG76069
 ID AAG76069 standard; Protein: 33 AA.
 XX
 AC AAG76069;
 XX
 DT 03-SEP-2001 (first entry)
 XX
 DE Human colon cancer antigen protein SEQ ID NO:6833.
 XX
 KW Human: colon cancer; colon cancer antigen; diagnosis; detection;
 KM colorectal carcinoma.
 XX
 OS Homo sapiens.
 XX
 PN WO200122920-A2.

XX
 PD 05-APR-2001.
 XX
 XX 28-SEP-2000; 2000WO-US26524.
 XX
 PR 29-SEP-1999; 99US-0157137.
 PR 03-NOV-1999; 99US-0163280.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Barash SC, Birse CE, Rosen CA;
 XX
 DR WPI: 2001-235357/24.
 DR N-PSDB: AAR35474.
 XX
 PT Nucleic acids encoding 427 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -
 XX
 PS Claim 11; Page 8285-8286; 9803pp; English.
 XX
 CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patients own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated Ps,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAB77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 XX
 SQ Sequence 33 AA:
 Query Match 1.6%; Score 7; DB 22; Length 33;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 8 LTLSTLS 14
 Db 22 ltlstls 28
 RESULT 15
 AAY79483
 ID AAY79483 standard; peptide: 50 AA.
 XX
 AC AAY79483;
 XX
 DT 01-AUG-2000 (first entry)
 XX
 DE Rat alpha tubulin variant peptide (aa314-363).
 XX
 KW Rat; alpha tubulin; mutant; mutein; variant;
 KW metabolic X syndrome; hypertension; stroke; diabetes;
 KW insulin resistance; obesity; dyslipidemia; anorectic;
 KW hypotensive; cerebroprotective; vasotropic; antidiabetic;
 KW antilipemic; gene therapy; spontaneously hypertensive; SHR.
 XX
 OS Rattus sp.
 XX
 FH Key
 FT Peptide
 Location/Qualifiers
 1..50
 /note="corresponds to residues 314-363 of

```

FT      Misc-difference 27      full-length protein"
FT      /note- "replaces Thr in wild-type"
XX
XX      WO200016918-A2.
XX
XX      06-APR-2000.
XX
XX      28-SEP-1999; 99WO-US22494.
XX
XX      28-SEP-1998; 98US-0161939.
XX
XX      (CURA-) CURAGEN CORP.
XX
XX      Shimkets RA;
XX
XX      WPI; 2000-303450/26.
XX
XX      Novel genes encoding e.g. CD36, SGLT2, and kynurenine aminotransferase,
PT      which are predictive and therapeutic for stroke, hypertension, diabetes
PT      and obesity -
XX
XX
XX      Claim 2; Fig 4G; 79pp; English.
XX
XX      This sequence represents amino acids 314-363 of alpha-tubulin
CC      from a spontaneously hypertensive rat (SHR). This region of
CC      alpha-tubulin carries a Ser for Thr amino acid substitution when
CC      compared to the corresponding region in control Wistar Kyoto (WKY)
CC      rats (see MAY79484). The alpha-tubulin gene is differentially
CC      expressed in SHR, SHR-stroke prone (SP) and WKY rats. It may be
CC      involved in a predisposition to vascular injury. The SHR rat
CC      is an animal model for human metabolic X syndrome. Genes encoding
CC      sodium dependent glucose cotransporter, kynurenine aminotransferase,
CC      CD36, aldolase A, atrial natriuretic factor, alpha-cardiac myosin
CC      and alpha-tubulin were identified as being potentially associated
CC      with hypertension, obesity and insulin resistance. These
CC      proteins, including isolated proteins comprising the present
CC      sequence, are used for treating, preventing and diagnosing
CC      ischemic and metabolic diseases and disorders, such as stroke,
CC      hypertension, diabetes and obesity, especially insulin resistivity,
CC      dyslipidemia and ischemic stroke (all claimed). Polynucleotides
CC      encoding them may also be used in gene therapy and antisense
CC      therapy protocols.
XX
XX      Sequence 50 AA:
SQ

```

Query Match 1.6%; Score 7; DB 21; Length 50;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      114 IKTIRST 120
        |||||
Db      22 Iktkrs1 28

```

Search completed: May 9, 2002, 12:54:59
Job time: 514 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2002, 12:55:42 ; Search time 36.87 seconds
(without alignments)
269.771 Million cell updates/sec

Title: US-09-787-083-6

Perfect score: 442
Sequence: 1 MKVSLSTLSTLSCFAI...YHNEATSPGVGLMDNMCL 442

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 212252 seqs, 22503292 residues

Word size : 0

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents_AA: *
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep: *
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/PCUS_COMB.pep: *
6: /cgn2_6/ptodata/2/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	1.6	16	1	US-07-942-245-98 Sequence 98, Appl
2	7	1.6	347	4	US-09-097-889-14 Sequence 14, Appl
3	7	1.6	552	1	US-08-231-7298-6 Sequence 6, Appl
4	7	1.6	900	4	US-08-890-865A-4 Sequence 4, Appl
5	7	1.6	934	1	US-08-215-805A-80 Sequence 80, Appl
6	7	1.6	1026	2	US-08-542-003-6 Sequence 6, Appl
7	7	1.6	1026	2	US-08-322-760A-6 Sequence 6, Appl
8	6	1.4	10	3	US-08-974-775-30 Sequence 30, Appl
9	6	1.4	11	3	US-08-974-775-29 Sequence 29, Appl
10	6	1.4	12	3	US-08-974-775-28 Sequence 28, Appl
11	6	1.4	13	3	US-08-974-775-8 Sequence 8, Appl
12	6	1.4	13	3	US-08-974-775-27 Sequence 27, Appl
13	6	1.4	14	3	US-08-974-775-9 Sequence 9, Appl
14	6	1.4	14	3	US-08-974-775-10 Sequence 10, Appl
15	6	1.4	14	3	US-08-974-775-12 Sequence 32, Appl
16	6	1.4	15	3	US-08-974-775-5 Sequence 5, Appl
17	6	1.4	16	3	US-08-974-775-4 Sequence 4, Appl
18	6	1.4	20	1	US-08-430-273-1 Sequence 1, Appl
19	6	1.4	21	1	US-07-918-953-13 Sequence 13, Appl
20	6	1.4	21	1	US-07-918-953-15 Sequence 15, Appl
21	6	1.4	21	1	US-08-212-696-1 Sequence 1, Appl
22	6	1.4	21	1	US-08-158-245-1 Sequence 1, Appl
23	6	1.4	21	1	US-08-081-661-13 Sequence 13, Appl
24	6	1.4	21	1	US-08-081-661-15 Sequence 15, Appl
25	6	1.4	21	1	US-08-233-617-1 Sequence 1, Appl
26	6	1.4	21	1	US-08-160-376A-1 Sequence 1, Appl
27	6	1.4	21	1	US-08-304-070-1 Sequence 1, Appl

28	6	1.4	21	1	US-08-285-661-1 Sequence 1, Appl
29	6	1.4	21	1	US-08-301-838-1 Sequence 1, Appl
30	6	1.4	21	1	US-08-389-487-4 Sequence 4, Appl
31	6	1.4	21	1	US-08-389-487-12 Sequence 12, Appl
32	6	1.4	21	1	US-08-507-124-2 Sequence 2, Appl
33	6	1.4	21	1	US-08-507-124-3 Sequence 3, Appl
34	6	1.4	21	1	US-08-342-931-1 Sequence 1, Appl
35	6	1.4	21	1	US-08-400-256-1 Sequence 1, Appl
36	6	1.4	21	2	US-08-508-664-9 Sequence 9, Appl
37	6	1.4	21	2	US-08-353-476-85 Sequence 85, Appl
38	6	1.4	21	2	US-08-353-476-87 Sequence 87, Appl
39	6	1.4	21	2	US-08-484-219-1 Sequence 1, Appl
40	6	1.4	21	2	US-08-979-587-1 Sequence 1, Appl
41	6	1.4	21	2	US-08-992-676-1 Sequence 6, Appl
42	6	1.4	21	2	US-08-992-676-6 Sequence 6, Appl
43	6	1.4	21	2	US-09-134-836-1 Sequence 1, Appl
44	6	1.4	21	2	US-09-134-836-7 Sequence 7, Appl
45	6	1.4	21	4	US-08-932-082-1 Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-07-942-245-98
Sequence 98, Application US/07942245
Patent No. 5639641
GENERAL INFORMATION:
APPLICANT: PEDERSEN, Jan T.
APPLICANT: SEARLE, Stephen M.J.
APPLICANT: REES, Anthony R.
APPLICANT: ROGUSKA, Michael A.
APPLICANT: GUILD, Brydon C.
TITLE OF INVENTION: SURFACE RESIDUE VENERING OF RODENT
TITLE OF INVENTION: ANTIBODIES
NUMBER OF SEQUENCES: 522
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughruue, Mion, Zim, Macpeak & Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: HP 9000/700 Workstation
OPERATING SYSTEM: UNIX
SOFTWARE: In house
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/942,245
FILING DATE: 09-SEP-1992
CLASSIFICATION: 530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 98:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-942-245-98

Query Match 1.6%; Score 7; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 7; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
QY 387 PRSGGA 393
|||||
Db 4 PRSGGA 10

RESULT 2
US-09-097-889-14
; Sequence 14, Application US/09097889
; Patent No. 6218117
; GENERAL INFORMATION:
; APPLICANT: Herrnstadt, Corrina
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Davis, Robert E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
; TITLE OF INVENTION: AGENTS THAT QUANTITATIVELY ALTER DETECTABLE
; TITLE OF INVENTION: EXTRAMITOCHONDRIAL DNA: MITOCHONDRIAL DNA RATIOS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/097,889
; APPLICATION NUMBER: US/09/097,889
; FILING DATE: 15-JUN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Roseman Ph.D., Stephen J.
; REGISTRATION NUMBER: 43,058
; REFERENCE/DOCKET NUMBER: 660088.417
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ. ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-09-097-889-14

Query Match 1.6%; Score 7; DB 4; Length 347;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 LITSLIS 14
|||||||
Db 155 LITSLIS 161

RESULT 3
US-08-231-729B-6
; Sequence 6, Application US/08231729B
; Patent No. 5618722
; GENERAL INFORMATION:
; APPLICANT: ZENNO, Shubel
; APPLICANT: SHIRAIISHI, Shinji
; APPLICANT: INOUE, Satoshi
; APPLICANT: SAIGO, Kaoru
; TITLE OF INVENTION: FIREFLY LUCIFERASE GENE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LEYDIG, VOIT & MAYER
; STREET: 700 Thirteenth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,729B
; FILING DATE: 20-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 119050/1993
; FILING DATE: 21-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rose, Herbert C.
; REGISTRATION NUMBER: 29846
; REFERENCE/DOCKET NUMBER: 60130/No. 5618722aka
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-737-6770
; TELEFAX: 202-737-6776
; INFORMATION FOR SEQ. ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 552 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-231-729B-6

Query Match 1.6%; Score 7; DB 1; Length 552;
Best Local Similarity 100.0%; Pred. No. 11e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 LSFOLDLR 181
|||||||
Db 184 LSFOLDLR 190

RESULT 4
US-08-890-865A-4
; Sequence 4, Application US/08890865A
; Patent No. 6307019
; GENERAL INFORMATION:
; APPLICANT: Constantin, Franklin
; APPLICANT: Zeng, Li
; TITLE OF INVENTION: AXIN GENE AND USES THEREOF
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/890,865A
; FILING DATE: 10-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/54249
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)278-0400
; TELEFAX: (212)391-0526
; INFORMATION FOR SEQ. ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 900 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-890-865A-4

Query Match 1.6%; Score 7; DB 4; Length 900;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 403 GKISGY 409
|||||
DB 268 GKISGY 274

RESULT 5
US-08-215-805A-80
Sequence 80, Application US/08215805A
Patent No. 5559008
GENERAL INFORMATION:
APPLICANT: Chang, Yung-Fu
TITLE OF INVENTION: LEUKOTOXIN GENE FROM PASTEURRELLA
TITLE OF INVENTION: SUI5
NUMBER OF SEQUENCES: 84
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: Rochester
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/215,805A
FILING DATE: 22-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Timian, Susan J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 19603/61 (D-1329A)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1636
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 934 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Pasteurella suis
STRAIN: 5943
IMMEDIATE SOURCE:
LIBRARY: P. suis DNA in Bacteriophage lambda-dash
CLONE: (lambda)yfc33-37
US-08-215-805A-80

Query Match 1.6%; Score 7; DB 1; Length 934;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 238 VKAEDL 244
|||||
DB 62 VKAEDL 68

RESULT 6
US-08-542-003-6

Sequence 6, Application US/08542003
Patent No. 5864013
GENERAL INFORMATION:
APPLICANT: Goldberg, Edward B.
TITLE OF INVENTION: MATERIALS FOR THE PRODUCTION OF
TITLE OF INVENTION: NANOMETER STRUCTURES AND USE THEREOF
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie and Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/542,003
FILING DATE: 13-OCT-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 8471-0005-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: 212-869-8664
TELEX: 66441 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1026 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Bacteriophage T4
IMMEDIATE SOURCE:
CLONE: p37 amino acid
US-08-542-003-6

Query Match 1.6%; Score 7; DB 2; Length 1026;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 128 QTIKGP 134
|||||
DB 861 QTIKGP 867

RESULT 7
US-08-322-760A-6
Sequence 6, Application US/08322760A
Patent No. 5877279
GENERAL INFORMATION:
APPLICANT: Goldberg, Edward B.
TITLE OF INVENTION: MATERIALS FOR THE PRODUCTION OF
TITLE OF INVENTION: NANOMETER STRUCTURES AND USE THEREOF
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie and Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/322,760A
FILING DATE: 13-OCT-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mastrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 8471-0003-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: 212-869-8864
TELEX: 66441 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1026 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Bacteriophage T4
IMMEDIATE SOURCE:
CLONE: p37 amino acid
US-08-322-760A-6

Query Match 1.6%; Score 7; DB 2; Length 1026;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 QTIKGR 134
|||||
Db 861 QTIKGR 867

RESULT 8
US-08-974-775-30
Sequence 30, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:
APPLICANT: Toback, F. Gary
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
STREET: Plaza Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775
FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid

STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-30

Query Match 1.4%; Score 6; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
|||||
Db 5 NHEATS 10

RESULT 9
US-08-974-775-29
Sequence 29, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:
APPLICANT: Toback, F. Gary
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
STREET: Plaza Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775
FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-29

Query Match 1.4%; Score 6; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 NHEATS 429
|||||
Db 5 NHEATS 10

RESULT 10
US-08-974-775-28
Sequence 28, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:

APPLICANT: Toback, F. Gary
APPLICANT: Walsh-Reitz, Margaret
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
STREET: Plaza Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775
FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-28

Query Match 1.4%; Score 6; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 424 NHEATS 429
Db 5 NHEATS 10
RESULT 11
US-08-974-775-8
Sequence 8, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:
APPLICANT: Toback, F. Gary
APPLICANT: Walsh-Reitz, Margaret
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
STREET: Plaza Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775

FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-8

Query Match 1.4%; Score 6; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 424 NHEATS 429
Db 8 NHEATS 13
RESULT 12
US-08-974-775-27
Sequence 27, Application US/08974775
Patent No. 6096706
GENERAL INFORMATION:
APPLICANT: Toback, F. Gary
APPLICANT: Walsh-Reitz, Margaret
TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
STREET: NBC Tower - Suite 3600, 455 N. Cityfront
STREET: Plaza Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60611-5599
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,775
FILING DATE: 20-NOV-1997
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Martin, Alice O.
REGISTRATION NUMBER: 35,601
REFERENCE/DOCKET NUMBER: 7814/27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-321-4200
TELEFAX: 312-321-4299
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-775-27
Query Match 1.4%; Score 6; DB 3; Length 13;

Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 424 NHEATS 429
|||||
DB 5 NHEATS 10

RESULT 13
US-08-974-775-9
; Sequence 9, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; STREET: Plaza Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ. ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-974-775-9

Query Match 1.4%; Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 424 NHEATS 429
|||||
DB 8 NHEATS 13

RESULT 14
US-08-974-775-10
; Sequence 10, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE

; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; STREET: Plaza Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ. ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-974-775-10

Query Match 1.4%; Score 6; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 424 NHEATS 429
|||||
DB 8 NHEATS 13

RESULT 15
US-08-974-775-32
; Sequence 32, Application US/08974775
; Patent No. 6096706
; GENERAL INFORMATION:
; APPLICANT: Toback, F. Gary
; APPLICANT: Walsh-Reitz, Margaret
; TITLE OF INVENTION: GROWTH-PROMOTING PROTEINS AND PEPTIDES
; TITLE OF INVENTION: FOR KIDNEY EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; STREET: Plaza Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,775
; FILING DATE: 20-NOV-1997
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 7814/27
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-321-4200
 TELEFAX: 312-321-4299
 INFORMATION FOR SEQ ID NO: 32:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 14 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FEATURE:
 NAME/KEY: modified-site
 LOCATION: 14
 OTHER INFORMATION: /product= "Ser or Tyr"
 US-08-974-775-32

Query Match 1.4%; Score 6; DB 3; Length 14;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 424 NHEATS 429
 |||||
 Db 8 NHEATS 13

Search completed: May 9, 2002, 12:55:42
 Job time: 552 sec

This Page Blank (uspto)